Understanding zebrafish pigment pattern formation using mathematical modelling

Owen, Jennifer

Award date:
2021

Awarding institution:
University of Bath

Understanding zebrafish pigment pattern formation using mathematical modelling

submitted by

Jennifer Owen

for the degree of *Doctor of Philosophy*

of the

University of Bath

Department of Biological Sciences

June 2021

Declaration of any previous Submission of the Work

The material presented here for examination for the award of a higher degree by research has not been incorporated into a submission for another degree.

. . . . . .

Jennifer Paige OWEN

Declaration of Authorship

I am the author of this thesis, and the work described therein was carried out by myself personally.

. . . . . .

Jennifer Paige OWEN

Contents

1 Abstract **1**

2 Introduction **2**
2.1 Overview . . . 2
2.2 Summary of original papers . . . 3
2.3 A note on the thesis format . . . 4

3 A literature review of zebrafish pigment pattern formation **5**
3.1 Overview . . . 5
3.2 Outline of the chapter . . . 5
3.3 Introduction . . . 7
3.4 Pigment producing cells in zebrafish . . . 7
3.4.1 Xanthophore . . . 7
3.4.2 Melanophore . . . 7
3.4.3 Iridophore . . . 7
3.5 Early larval pigment pattern . . . 8
3.5.1 Embryonic/larval pigment cells and adult pigment stem cells . . . 8
3.6 The composition of the adult striped pattern . . . 9
3.6.1 Adult zebrafish pigment stem cells and their niche . . . 9
3.7 Adult zebrafish pattern development . . . 10
3.7.1 Developmental staging . . . 10
3.7.2 Adult pattern development . . . 11
3.7.3 The initial iridophore interstripe provides a morphological landmark for pattern development . . . 12
3.8 All three pigment cell-types are required to generate the adult pigment pattern . . . 12
3.8.1 No iridophores (*ltk/shady*) . . . 13
3.8.2 No melanophores (*mitfa/nacre*) . . . 13
3.8.3 No xanthophores (*csf1a/c-fms/fms/panther/pfeffer/salz*) . . . 14
3.8.4 Double mutants of *shd*, *nac* and *pfe* . . . 14
3.8.5 Two missing pigment producing cells (*nac;pfe*, *shd;pfe*, *shd;nac*) 16
3.9 Regulation of cellular processes . . . 16
3.9.1 Thyroid hormone . . . 16
3.10 Reduced iridophores in the adult . . . 16
3.10.1 *alkal2a; alkal2b* . . . 17
3.10.2 *ednrb1/ednrb1a/ rse* . . . 17
3.10.3 Edn3 . . . 18
3.10.4 Mpv17/*tra* . . . 18
3.11 Genes affecting the pigment cell environment . . . 18
3.11.1 *bnc2* . . . 19
3.11.2 Ece2/ *kar* . . . 19
3.11.3 Somatolactin . . . 19
3.12 Genes required for early metamorphic melanophores . . . 20
3.12.1 *kit/spa* and *kitla/slk* . . . 20
3.13 Genes required for metamorphic chromatophore development . . . 20
3.13.1 *tuba8l3/puma* . . . 21

3.13.2 Erbb3b/*picasso* . . . 21
3.14 Genes relating to cell shape . . . 21
3.14.1 *bace2/wanderlust* . . . 22
3.15 Cell communication during pattern development . . . 22
3.15.1 Cell-studies into cell communication . . . 22
3.15.2 Ablation studies to assess cell-cell communication . . . 24
3.15.3 Genetic studies into cell communication . . . 24
3.15.4 Potential molecular mechanisms mediating pigment cell interactions . . . 28
3.16 Summary of mutations that cause defective patterns . . . 29
3.17 Dorso-ventral countershading . . . 38
3.17.1 Scale development . . . 38
3.17.2 Genes relating to dorso-ventral countershading (DVC) Agouti signaling . . . 38
3.17.3 Melanocortin receptors . . . 39
3.17.4 Suggested molecular mechanisms . . . 39
3.18 Conclusions . . . 40
3.19 Chapter Conclusions . . . 41

4 A literature review of the mathematical modelling of zebrafish pattern formation 42
4.1 Overview . . . 42
4.2 Outline . . . 42
4.3 Introduction . . . 43
4.4 Component-level models . . . 43
4.4.1 Heterotypic repulsion and homotypic attraction between melanophores and xanthophore can generate stripes . . . 44
4.4.2 Community effect may be a viable patterning mechanism . . . 45
4.4.3 Chase and run between melanophores and xanthophores is not sufficient for pattern formation . . . 45
4.4.4 Summary of component-level models . . . 46
4.5 Comprehensive models . . . 46
4.5.1 Continuous models . . . 47
4.5.2 Discrete models . . . 51
4.5.3 Summary of comprehensive models . . . 56
4.6 Conclusion . . . 56
4.7 Chapter Conclusion . . . 61

5 Pair correlation functions for identifying spatial correlation in discrete domains 62
5.1 Overview . . . 62
5.2 Outline of the paper . . . 62
5.3 Conclusions . . . 89

6 A quantitative modelling approach to zebrafish pigment pattern formation 90
6.1 Overview . . . 90
6.2 Outline of the chapter . . . 90
6.3 Conclusions . . . 155

7 A new anaesthetic protocol for metamorphic zebrafish 157
7.1 Overview . . . 157
7.2 Outline of the paper . . . 157
7.3 Abstract . . . 159
7.4 Introduction . . . 159
7.5 Materials and Methods . . . 161
7.5.1 Ethics Statement . . . 161
7.5.2 Facility conditions . . . 161
7.5.3 Schedule 1 killing . . . 162
7.5.4 Anaesthetic preparation . . . 162
7.5.5 Experiment A . . . 162
7.5.6 Experiment B . . . 163
7.5.7 Experiment C . . . 165
7.5.8 Experiment D . . . 166
7.5.9 Statistics . . . 166
7.6 Results . . . 167
7.6.1 Experiment A - longer duration anaesthesia with recovery cannot be readily obtained using MS-222 for juveniles within a sensitive growth range . . . 167
7.6.2 Length, rather than age, is a better predictor of successful anaesthesia by MS-222 . . . 167
7.6.3 Pre-metamorphic (<6mm SL) and J+ (>16mm SL) fish can be anaesthetised and recovered successfully with MS-222 . . . 169
7.6.4 Metamorphic (8 mm SL - 16 mm SL) fish react unpredictably to MS-222 . . . 169
7.7 Experiment B - Alternative anaesthetics and maintenance methods have potential for repeated anaesthesia for juvenile fish . . . 169
7.7.1 Anaesthetic choice affects times taken to induce, lose touch response and recover post anaesthetic . . . 171
7.7.2 Maintenance of anaesthesia via cotton pad, as opposed to immersion, improves recovery from anaesthesia . . . 171
7.8 Experiment C - Towards a reliable methodology for repeated anaesthesia for juvenile fish . . . 172
7.8.1 Full recovery was observed for protocols 1 and 4 only . . . 173
7.8.2 Repeated anaesthesia did not impact on fish growth . . . 173
7.8.3 Protocols 4 and 6 achieve anaesthesia most rapidly . . . 174
7.8.4 Protocol 4 suppresses breathing rate. . . . 176
7.8.5 Protocol 4 is most suitable method based on time taken to recover 176
7.8.6 Maintenance of anaesthesia via cotton pad, as opposed to immersion, improves recovery and time taken to recover from anaesthesia for propofol/lidocaine mix . . . 177
7.8.7 Protocol 4 is the most suitable protocol for repeated anaesthesia. 178
7.9 Experiment D - Protocol 4 is effective for frequent, repeated anaesthesia with recovery throughout the susceptible juvenile period . . . 178
7.9.1 Growth of the fish was not retarded by the repeated anaesthesia at a higher frequency . . . 178
7.9.2 Protocol 4 remained an efficient and effective method for inducing anaesthesia with recovery when used frequently. . . . 180
7.9.3 Protocol 4 is a suitable protocol for repeated, frequent anaesthesia. 180
7.10 Discussion . . . 180

7.11 Conclusions . . . 201

8 Differential growth is a critical determinant of zebrafish pigment pattern formation 202
8.1 Overview . . . 202
8.2 Outline of the chapter . . . 202
8.3 Summary paragraph . . . 204
8.4 Introduction . . . 204
8.5 Results . . . 205
8.5.1 Initial iridophore position is not predictive of the orientation of pattern . . . 205
8.5.2 Changing the rate of growth markedly changes patterning . . . 205
8.5.3 Impact of biased growth on multiple features of zebrafish pigment pattern . . . 207
8.5.4 Biased growth explains asymmetry in missing cell-type and *rose* mutants . . . 211
8.5.5 Biased growth may contribute to the DV asymmetry of Danio nigrofasciatus . . . 211
8.6 Discussion . . . 213
8.7 Methods . . . 214
8.7.1 Ethics Statement . . . 214
8.7.2 Repeated imaging of juvenile fish . . . 214
8.7.3 2-phenoxyethanol . . . 214
8.7.4 Mathematical model of zebrafish pigment pattern formation . . . 214
8.7.5 Implementation of *Danio nigrofasciatus* . . . 217
8.7.6 Dorsal -ventral relocation of melanophores . . . 217
8.7.7 Pair correlation function . . . 217
8.8 Statistics . . . 217
8.8.1 Linear regression . . . 217
8.9 Conclusions . . . 229

9 Final conclusions and outlook 230
9.1 Future Research . . . 231
9.1.1 Derivations of continuum models of zebrafish pattern formation . . . 231
9.1.2 Gene analysis . . . 231
9.1.3 Extending the model beyond juvenilehood . . . 232
9.1.4 Dorso-ventral countershading . . . 232
9.1.5 Beyond zebrafish . . . 233

Appendices 234

Appendix A Dorso-ventral countershading: A preliminary study 235
A.1 Materials and Methods . . . 236
A.1.1 Fish . . . 236
A.1.2 Imaging . . . 237
A.2 Results . . . 237
A.2.1 Dorso-ventral countershading is related to cell composition in the dorsal and ventral scales . . . 237
A.2.2 Melanocyte and xanthophore counts vary dorso-ventrally across all fish types investigated . . . 237

A.2.3 Iridophores on the scales appear in two different shapes 239
A.3 Predictions 241
A.3.1 An increase in *asip1* is associated with a decrease in melanocyte production (when all receptors are present) 241
A.3.2 *asip1* is not the only ligand responsible for the dorso-ventral gradient of melanocyte production 241
A.3.3 In the absence of Mc1r, *asip1* is a potential agonist for another melanocortin receptor 241
A.4 Conclusion 241

Chapter 1

Abstract

Animal pigment patterns are as diverse as they are beautiful. Characterised by their striking blue and golden stripes along the body tails and fins, the adult *Danio rerio* (zebrafish) has become a major model organism for the study of pigment patterns. The stripes are comprised of three different pigment producing cell types which self-organise over a period of three weeks into the adult stripe pattern. Over the past two decades, a multitude of observational, cell and genetic studies have determined numerous mechanisms involved in this self-organisation, yet, it has still been unclear whether or not these rules were sufficient to generate the pattern. To aid analysis of complex systems such as this, mathematical modelling has become a powerful assistive tool. By providing a theoretical framework, a mathematical model allows us to clearly articulate our hypotheses in a quantitative manner and then test these hypotheses.

The goal of this thesis is to develop a mathematical model of zebrafish stripe formation which will be used both within this thesis and in the future, in conjunction with biological experimentation to elucidate the key mechanisms underlying zebrafish stripe formation. We will start by introducing the known mechanisms as determined from biological experiments, as well as introducing the findings and limitations of previous mathematical models of zebrafish stripe formation. Based on this knowledge, we will build our own comprehensive mathematical model of zebrafish stripe formation incorporating all the up-to-date biological information. As we will demonstrate, not only does this model quantitatively and qualitatively reproduce the wild-type striped pattern, but it can also reproduce many of the mutant patterns. Thus we show that the rules determined from the biology are, to at least a good approximation, sufficient for producing the striped patterns. Having validated our approach, we next use the model as a predictive tool to guide further investigation. For example, in the last chapter of the thesis, we will use the model alongside biological experimentation to explore the importance of growth during pattern metamorphosis in the context of both *Danio rerio* and its close relatives. Along the journey of this interdisciplinary study, we will develop new techniques across mathematics and biology that will aid this investigation. For example, in order to validate our mathematical models ability to generate wild-type zebrafish stripes, as well as other mutant patterns, we will develop a pair correlation function: a statistical tool used for quantifying spatial correlation. In order to compare our theoretical model with the biology, we will refine a method for anaesthetising and subsequently imaging juvenile fish during pattern development that minimises suffering. Therefore the outcomes of this thesis not only aids our understanding of pattern formation, but provides a framework for future investigations of its kind.

Chapter 2

Introduction

2.1. Overview

Patterns are abundant in nature. Examples range from Fibonacci shells [1] to the spiral arrangement of leaf patterns [2]. The study of pattern formation has drawn attention from both mathematical and biological communities alike due to its often inherent simplicity resulting in complex and beautiful structures.

Among all the examples of patterns observed in nature, those which are observed on the body of animals such as stripes and spots are of particular interest. Coloured animal pigment patterns (such as the spots of the leopard and the stripes of tigers), are extraordinarily diverse and serve a crucial role in determining mate choice, distinguishing between individuals, environmental adaptation and predator avoidance [3, 4].

Within pattern formation studies, zebrafish (*Danio rerio*) have become a model organism [5, 6, 7, 8, 9, 10]. Wild-type (WT) adult zebrafish display 4 to 5 stripes interspersed by 4 yellow interstripes that form during metamorphosis as mosaics of yellow xanthophores, silvery or blue iridophores and black melanophores in the hypodermis [4, 11]. A major reason that zebrafish have become a model organism is that the pigment pattern formation process is well-documented [12]. The process is well-documented for the following reasons: zebrafish are amenable to many observational studies (development takes place outside the mother, allowing direct inspection at all stages); they are small in size and transparent, allowing repeated imaging; and are accessible for genetic and mutational analyses providing unique opportunities to access the cellular and molecular background of the developmental process [5, 6, 7, 8, 9, 10].

Despite this, due to the complexity of the pattern development process, there are still many unanswered questions. First and foremost, it is not clear if the mechanisms identified in the literature are sufficient to explain the pattern forming process. Answering this question gives a more complete understanding of the main mechanisms of zebrafish stripe formation. With this understanding, more detailed questions can be addressed. For example, mutations in genes related to cell-cell communication can cause dramatic changes to the pattern such as the production of spots instead of stripes, yet their exact functionality is unknown. Also, the phenomena in which the pattern appears asymmetric over the dorso-ventral axis is currectly unexplained and is unexpected due to the supposed symmetric nature of all known underlying mechanisms.

Mathematical modelling plays a central role in addressing such questions by providing a tool to test hypotheses and draw comparisons with data. In particular, based upon a finite set of rules indicated from the literature, comprehensive (or more simplistic component-type) models can be developed, tested and then in an iterative process fine-tuned based on the results. Mathematical modelling of zebrafish stripe formation and pattern formation more generally has became more and more popular in the past decade [13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26].

The aim of this research is to improve the understanding of zebrafish stripe formation using a comprehensive mathematical modelling technique. The thesis can be split into two parts.

The first part of the thesis is concentrated on building a robust and valid model that

is sufficient for determining WT and mutant zebrafish patterns. To do this, we build a comprehensive on-lattice volume exclusion mathematical model of zebrafish stripe development comprising all of the known literature relating to the stripe development. To validate the model we generate a new method for quantifying spatial correlation in discrete domains which we use to quantify the sizes of simulated stripes and spots and compare it to real data. Satisfied that the model is valid, we use the model to deduce that the cell-cell interactions known in the literature are sufficient for generating WT and mutant patterns as well as using it to investigate the functionality of some genes related to pattern formation.

The second part of the thesis is focused on using the model developed in the first part alongside experimental results to investigate dorso-ventral asymmetry observed in *Danio* patterns. Based on the prediction that asymmetry is caused by biased growth ventrally, we adapt the model to incorporate different types of growth bias. Simultaneously we develop a method to repeatedly image juvenile zebrafish during pattern development which we use to parameterise the bias of the growth in the model. Using the parameterised model we demonstrate that biased growth can explain the asymmetry of pattern formation as well as other previously unexplained phenomena [27].

The work in this thesis has led to a book chapter and four original papers (two published [28, 29], one accepted for publication and one in preparation). The review chapter is presented in Chapter 3 and provides a detailed investigation into all experimental results relating zebrafish pattern formation. In Chapter 4 we perform a subsequent review of all mathematical models relating to zebrafish stripe pattern formation. Chapters 3 and 4 do not include original results, but are provided in this thesis as a general introduction to the field. The rest of the thesis is presented following the Alternative Format or Thesis by Publications style. That is, each of the remaining chapters (Chapters 5 - 8) comprises one of the published manuscripts. Each of these chapters are prefaced with a short introduction and finish with a conclusion which helps to place the corresponding chapter into the context of the thesis. In Chapter 9, we conclude by discussing the results of the thesis from a broader perspective along with suggestions of future work.

2.2. Summary of original papers

Chapter 5- Pair correlation functions for identifying spatial correlation in discrete domains
In this chapter we present our work identifying and quantifying spatial correlation in multi-agent systems. We develop a set of statistical functions, specifically designed for discrete domains, which determine whether the positions of agents are positively or negatively correlated across a range of distances. These functions are used in Chapter 6 to quantify stripes or spots in our on-lattice mathematical model of zebrafish pattern development and subsequently validate our modelling approach.

Chapter 6- A quantitative modelling approach to zebrafish pigment pattern formation
In this chapter we present a comprehensive mathematical model that incorporates all of the interactions described in Chapter 3 into a stochastic on-lattice model that can replicate both WT and mutant patterns. This comprehensive model of zebrafish pattern formation can then be used to investigate individual aspects of pattern development, such as the functions of genes related to pigment pattern formation as well as those presented in Chapter 8.

Chapter 7- A new anaesthetic protocol for metamorphic zebrafish
In this chapter, motivated by the prospect of generating parameter data extracted from imaging real zebrafish during development, we present a refined method for repeatedly anaesthetising and recovering fish sized between 8 mm SL - 16 mm SL. This method is used in Chapter 8, where we repeatedly image juvenile fish in order to record the position of an interstripe during pattern development and use this information to parameterise the model accordingly.

Chapter 8- Growth explains key aspects of vertebrate development
In this chapter we present a detailed study into the role of growth during zebrafish pattern formation. Using the model developed in Chapter 6 we provide evidence that the location of domain expansion in zebrafish is biased ventrally, showing that this feature, when incorporated into our model, can explain many previously unexplained features of WT development as well as the asymmetry observed in many mutants. Furthermore we show that growth plays an important role more generally in determining the type of patterns formed.

2.3. A note on the thesis format

Chapter 3 contains a review chapter published in Springer Book entitled: 'Pigments, Pigment Cells, Pigment Patterns'. Chapter 5 contains a paper published in Physical Review E. Chapter 6 contains a paper published in eLife. Chapter 7 contains a paper published in PLOS ONE. Chapter 8 contains a preprint of a paper that will be submitted shortly. All the published material in this thesis is presented in the original format of the publication. The paper in Chapter 8 is formatted to match the style of the thesis.

Chapter 3

A literature review of zebrafish pigment pattern formation

3.1. Overview

The aim of this chapter is to give an overview of zebrafish pattern formation, with particular emphasis on the known mechanisms involved in the process. These findings will provide the context for the review of mathematical models of zebrafish stripe formation given in Chapter 4 as well as the biological foundations with which we will use to build our own mathematical model of zebrafish stripe formation in Chapter 6. This chapter is adapted from the book chapter written by myself (Jennifer Owen), Christian Yates, Robert N. Kelsh, due to appear in 'Pigments, Pigment Cells, Pigment Patterns' published by Springer.

In this book chapter we begin by describing the composition of the adult zebrafish striped pattern as well as the the generation of the adult pattern. We then provide a detailed summary of all known genetic, cell biology and ablation studies that inform our knowledge of the mechanisms involved with pattern formation.

In the context of the thesis, the role of this chapter is to provide the relevant background for building a biologically-founded mathematical model of zebrafish pattern formation. In particular, all of the descriptions of cell proliferation, death, migration and cell-cell interactions that are outlined here are incorporated into our model in Chapter 4.

3.2. Outline of the chapter

In Section 3.3, we give a brief introduction to zebrafish stripe formation. In Section 3.4 we introduce the three pigment producing cell types involved in zebrafish pattern formation. In Section 3.5, we describe the early larval pattern from which the adult pattern is transformed. In Section 3.7 we describe how the adult pattern develops. In Sections 3.8 - 3.16 we outline all known genetic, cell and ablation studies relating to stripe formation. Finally, Section 3.19 contains a short concluding discussion.

Appendix B: Statement of Authorship

This declaration concerns the article entitled:			
Zebrafish Pigment Pattern Formation			
Publication status (tick one) **Draft manuscript** ☐ **Submitted** ☐ **In review** ☐ **Accepted** ☐ **Published** X			
Publication details (reference)	Book: Pigments, Pigment Cells and Pigment Patterns, Chapter 8 Authors: Jennifer P. Owen, Christian A. Yates and Robert N. Kelsh		
Copyright status (tick the appropriate statement) I hold the copyright for this material ☐ Copyright is retained by the publisher, but I have been given permission to replicate the material here X			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	All the content provided in this chapter was written by the author of the thesis (100%). All authors contributed equally to the presentation of the content and to the final edit (33 %).		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed	Jennifer Owen	**Date**	07/06/2021

3.3. Introduction

Pigment patterns in adult fish show astonishing diversity, reflecting their likely roles in camouflage, kin- and mate-selection, and display. Adult pigment pattern formation has only been explored extensively in the zebrafish, *Danio rerio*, and here we summarise the experimetnal studies that help to define a myriad of cell-cell interactions that result in this pattern. We begin by summarising zebrafish stripe morphology and the general process whereby the adult pattern forms. We then provide a detailed summary of the genetic, cell biology and ablation studies that inform our knowledge of the biological bases for stripe formation.

3.4. Pigment producing cells in zebrafish

Zebrafish pattern formation is caused by the self-organisation of three kinds of pigment producing cells: xanthophores, melanophores and iridophores. The composition of these cells and their pigment producing properties are described below.

3.4.1 Xanthophore

Xanthophores utilise pteridines and carotenoids as their pigmentary compounds. Biochemical differences in the balance of their carotenoid composition appear to drive their specific colour. Since these colours range in a continuum through yellow to red, we consider these two chromatophores as variants of a single cell-type, which we will refer to as xanthophores in this chapter.

3.4.2 Melanophore

Melanophores in zebrafish produce eumelanin. They are sometimes called melanocytes, since their developmental genetics shows strong homology to that of mammalian melanocytes. However, fish melanophores show two distinctive features compared to mammalian melanocytes:

1. they do not secrete melanin-containing melanosomes; and
2. melanosomes are actively moved radially along the microtubules, causing them to either aggregate in the perikaryon or to disperse throughout the cytoplasm (aggregation-dispersion response) in response to environmental and/or nervous cues.

3.4.3 Iridophore

Iridophores are defined by Transmission Electron Microscopy (TEM) through having thin, membrane-bound guanine crystals, called reflecting platelets, in the cytoplasm. Whether they look blue, silvery or whitish depends on the arrangement of the reflecting platelets, e.g. silver hue is dependent upon platelets being organized as uniformly-spaced stacks. Iridophores contain 'reflecting platelets', stacks of long hexagonal guanine crystals, while leucophores contain leucosomes, which are rounded, when observed by TEM.

In taxonomic terms, iridophores are widespread, being found in teleosts, as well as in amphibians and reptiles. Iridophores in the body pattern appear in two forms: L-type iridophores which form a homogeneous silvery sheet below the melanophores and S-type iridophores which form above the melanophores and xanthophores.

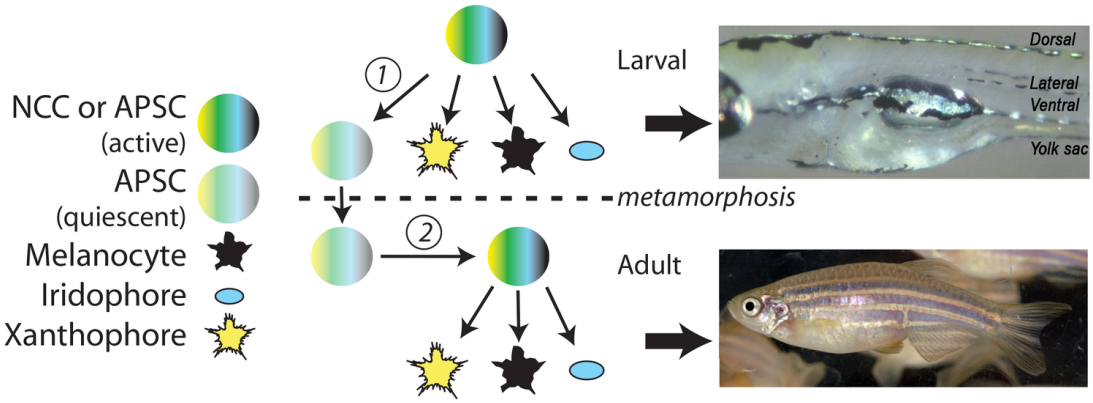


Figure 3-1: ***Differentiating melanophores first appear at approximately 24 hours postfertilization (hpf) in the dorsolateral trunk and head [30, 31].*** *These cells begin to synthesize melanin even as they are migrating, although they remain only partially-pigmented until they settle in their final locations. Xanthophores containing pteridine pigments are first evident at approximately 42 hpf as pale yellow coloration on the dorsal aspect of the head. Around this same time iridophores are first observable around the choroid of the eye and later along the dorsal midline of the trunk and tail. Production of each cell-type continues and, by 6 days postfertilization, the embryonic pigment pattern is established: four horizontal melanophore stripes, three with associated iridophores, and xanthophores throughout the skin dorsal to the gut along the length of the body. This larval pattern remains unchanged until the start of adult pattern metamophosis at which point metamorphic melanophores, xanthophores and iridophores start to differentiate on the skin. The pattern forms sequentially, starting with initial iridophore interstripe, facilitated by cell-cell interactions until the final adult pattern forms.*

3.5. Early larval pigment pattern

Zebrafish have two pigment patterns, an early larval pattern which is developed during the embryonic stages and an adult pattern. The larval pattern consists of four distinct dark stripes of early larval melanophores, as well as larval xanthophores which cover the flanks just under the skin [32, 30, 27]. The early larval stripes are located as follows: Dorsal Stripe, along the dorsal apex of the myotomes and continuing along the dorsal surface of the brain; Ventral Stripe, at the ventral apex of the myotomes and anteriorly under the gut. Lateral Stripe, along the outer half of the horizontal myoseptum. Yolk Sac Stripe, medially along the belly from under the heart to the anus. Many of these early larval melanophores are lost during metamorphosis, but retained melanophores can be distinguished from metamorphic melanophores by their brown (not black) colour. Iridophores are distributed widely in three of the stripes (Dorsal, Ventral and Yolk Sac), usually as scattered cells, but sometimes in contiguous sheets (Lateral Patches of the Ventral Stripe, and in a continuous sheet outside the melanophores of the Yolk Sac Stripe.

This basic larval pattern is replaced during metamorphosis with an adult pattern comprised of 4-5 dark blue stripes, interspersed with 4 golden interstripes across the body tail and fins. The adult pattern formation is the development period we are interested in and will be the focus of this review.

3.5.1 Embryonic/larval pigment cells and adult pigment stem cells

Pigment cells are generated during two distinct stages of fish development, resulting in two distinct patterns. Production of embryonic/larval cells directly from neural crest cells (NCCs) generates the early larval pattern, whilst generation of new cells during metamorphosis from adult pigment stem cells (APSCs) generates the adult pattern

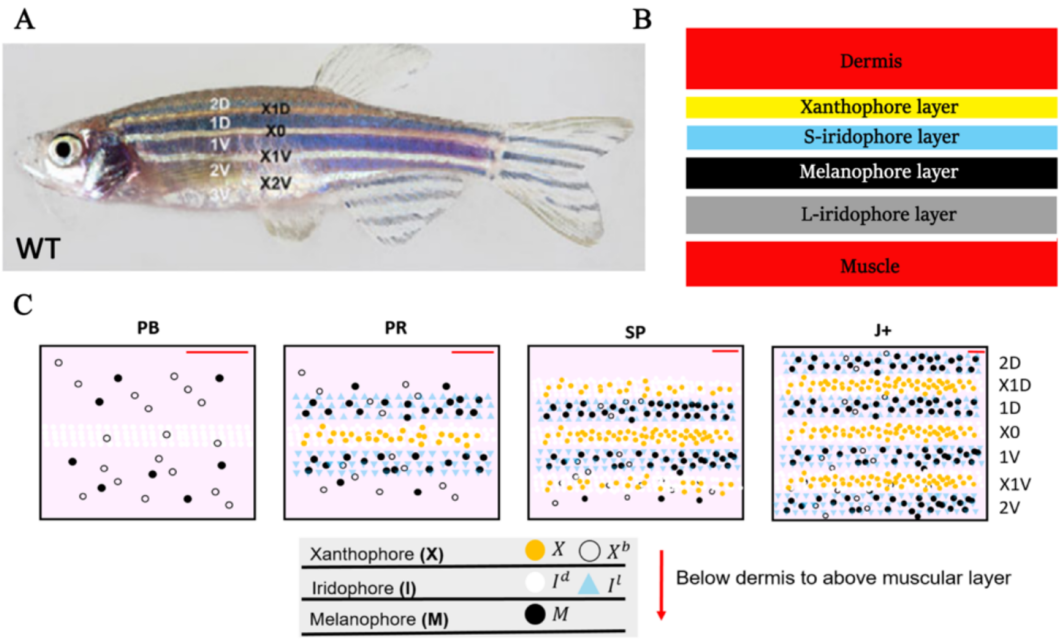


Figure 3-2: ***WT zebrafish stripe composition and development*** *(A) An adult wild type (WT) fish. Stripes and interstripes are labelled according to their order of temporal appearance and position on the body. X0 is the first interstripe to appear. 1D and 1V (D - Dorsal, V- Ventral) are the first two stripes to appear. X1D and X1V are the next two interstripes to appear and so on. (B) Summary of pigment cell distribution in adult zebrafish. The cells in the xanthophore, S-iridophore, melanophore and L-iridophore layers consist of xanthophores and xanthoblasts, S-iridophores, melanophores and L-iridophores, respectively. (C) Schematic of WT patterns on the body of zebrafish. Stages PB, PR, SP, J+ correspond to developmental stages described in Table 3.1. Patterns form sequentially outward from the central X0 interstripe. Images reproduced from Owen et al [28] and licensed under CC-BY 3.0(http://creativecommons.org/licenses/by/3.0); published by The Company of Biologists Ltd.*

(Fig. 3.4.3).

3.6. The composition of the adult striped pattern

The adult pattern comprises of the superimposition of three layers of pigment cell-types [33]. Starting from the deepest layer of the hypodermis just above the muscle and moving to the dermis, the dark stripes consist of consecutive layers of L-iridophores, melanophores and loose S-iridophores (Fig. 3-2 (A)). Similarly, light interstripes are made up of layers of S-iridophores and xanthophores. It is the iridescent blue colouration of the loose S-iridophores lying over black melanophores that gives the blue-black colour to the dark stripes, whilst the yellow xanthophores lying over the iridescent silver dense S-iridophores give the yellow appearance to the pale interstripes.

3.6.1 Adult zebrafish pigment stem cells and their niche

Whereas embryonic/larval pigment cells are directly derived from neural crest cells, most adult pigment cells are not generated from migrating neural crest cells but rather from a reserved population of postembryonic neural crest-derived progenitors (Fig. 3.4.3). In zebrafish, melanophores and iridophores are newly generated in the adult: embryonic/larval cells are replaced with postembryonic stem cells during metamorphosis. In contrast, adult xanthophores are derived from two distinct sources of pro-

Stage	Acronymn	Size (SSL)	∼Age (dpf)
Pelvic Bud	PB	7.2	21
Pelvic Ray	PR	8.6	30
Squamation onset Posterior	SP	9.6	39
Squamation through Anterior	SA	10.4	44
Juvenile	J	11	51
Juvenile+	J+	13.5	70
Juvenile++	J++	16	90

Table 3.1: *Developmental stages of development as determined by (Parichy et al., 2009) [31].*

genitors: one is postembryonic multipotent APSCs and the other is embryonic/larval xanthophores which undergo a remarkable process of dedifferentiation, proliferation and then xanthophore differentiation. The location of APSCs in the fish is only partially resolved. There is strong evidence that at least some are located in the peripheral nervous system, most clearly shown for the dorsal root ganglia (DRGs), but likely also to include the spinal nerves [31, 34]. In the DRGs, an mitfa:GFP reporter transgene labels a small subset of ganglial cells. APSCs were initially shown indirectly to be set aside during a defined, embryonic time-window, in a manner that was ErbB-dependent, with adult pigment pattern being severely disrupted in adult *erbb3b* mutants. Consistent with this, in the absence of ErbB signaling formation of DRGs is defective and mitfa:GFP labelled cells are absent. Melanophores are delivered to the adult skin via the peripheral nerves, and do so in a manner that initially places them in widely-scattered locations throughout the flank. Adult iridophores are also derived from similar APSCs in the DRGs, and are largely lost in *erbb3b/picasso* mutants. Iridoblasts appear to migrate through the horizontal myoseptum where they continue to migrate and spread after arrival in the skin. Iridoblasts proliferate dramatically in the skin, in contrast to melanoblasts, which are less proliferative after fate-specification. Melanoblasts proliferate while migrating along the peripheral nerve. Lineage tracing has shown that these APSCs in the DRGs remain broadly multipotent and give rise to some xanthophores, neurons and glia, in addition to the metamorphic melanophores and iridophores. Understanding pigment pattern formation has come from various studies, but key amongst these have been careful analysis of the pigment pattern defects in a series of adult pigment pattern mutants, defined as those where the distribution (as opposed to the colouration) of pigment cells in the adult is disrupted. Here we will discuss these, dividing them into classes by phenotype.

3.7. Adult zebrafish pattern development

3.7.1 Developmental staging

During adult metamorphosis the zebrafish grows significantly [31, 35] Developmental staging of zebrafish is typically measured using standard length (SL) or standardized standard length (SSL). Standard length is defined as the distance from the tip of the snout to the caudal peduncle. Standard length is used as a proxy for developmental stage in zebrafish as it can explain significantly more developmental variance than days post fertilization (dpf) and hence is preferred over dpf as an indicator of developmental progress [31]. The standardised standard length is the predicted standard length based on developmental progress, *i.e.* normalized for individual variation or strain differences. A list of developmental stages and their approximate times are given in Table 3.1.

3.7.2 Adult pattern development

Transformation of the early larval pattern to the adult pigment pattern starts at the onset of metamorphosis. Thyroid hormone acts as a trigger for metamorphosis and also regulates xanthophore proliferation and differentiation [36]. Our understanding of the molecular basis for pattern formation is very limited, although identification of a diverse set of mutants affecting the process provides clues to drive future understanding. Our understanding of the cell-cell interactions that drive patterning is much better, and we will focus on this in the following summary. The start of adult stripe development is marked by the differentiation of metamorphic S-iridophores around 7.2mm SL (3wpf) along the horizontal myoseptum, which serves as a morphological prepattern determining the orientation of the adult stripes [31, 35]. From the horizontal myoseptum, the iridophores migrate and proliferate dramatically (doubling time of 3-4 days [37], performing patterned shape transitions between dense form (interstripe regions) and loose form (stripe regions) regulated by tight junction proteins [38, 39]). Tight junction protein 1a, for example, regulates the transition from dense to loose and in fish where this is disrupted, dense iridophores extend to form a coherent sheet over the entire stripe region [7]. Iridophores move without directionality. However, the availability of space along the dorsal-ventral axis coupled with competition from other iridophores along the anteroposterior axis appear to orient the population of migrating iridophores bidirectionally along the dorso-ventral axis [39]. By local proliferation, iridophores spread to typically 2-4 segments (87% of 163 labelled iridophore clones by Singh et al spanned at least two interstripes [39]). Not all iridophores in the interstripe of one segment are clonally related, indicating a mutual mixing of iridophores from neighbouring segments. However, most iridophore clones span at least two interstripes, including the first one, suggesting that most of the later arising iridophores are clonally derived from iridophores of the initial interstripe X0 [39]. A combination of extensive migration, shape transition and proliferation produce the striping observed. Most metamorphic xanthophores arise from larval xanthophores [40]. Larval xanthophores suddenly dedifferentiate in response to the release of thyroid hormone, before dividing rapidly [38]. Iridophores express Colony stimulating factor-1 (Csf1), which promotes local differentiation of some xanthophores in the interstripes [41]. Many larval xanthophores do not differentiate and instead remain invisible as unpigmented xanthoblasts in the prospective stripes [40]. Xanthophores migrate slowly, at a rate approximately 3 times slower than melanophores (approximately 33μm per week) *in vitro* [42], and are generally observed to be slower than melanophores *in vivo*. Long-term time lapse imaging of labelled xanthophore clusters show that differentiated xanthophores maintain their relative positions during stripe formation. Though differentiated xanthophores proliferate continuously (a doubling time of about 1 week [38] and interact with newly differentiated cell-types during stripe morphogenesis, there is no global reorganisation of xanthophores during stripe formation [38]). Larval melanophores play only a minor role in *Danio rerio* adult pattern development, though they are more important in other species of *Danio*. Residual larval melanophores occupying the ventral and dorsal position persist, shifting slightly dorsally and ventrally during pattern metamorphosis but otherwise remain unchanged [27]. A minority of the lateral early larval melanophores contribute to adult primary stripe 1D (Fig. 3-2), most die in the interstripe following its development of dense iridophores and xanthophores [27, 43]. The adult melanophore population is instead dominated by the appearance of metamorphic melanophores. During stripe development, progenitors of metamorphic melanophores migrate to the skin and begin to differentiate widely over the flank, though most melanophores differentiate in prospec-

tive stripe regions. Melanophores migrate slowly (a rate of approximately 100μm per week in vivo) [43]. Furthermore, melanoblasts and melanophores rarely proliferate (in a study by Singh *et al* of pattern metamorphosis, it was observed that only 4 out of 25 melanoblasts and 2 out of 47 melanophores divided once in 9 days [39]. As the stripes form, melanophores present in prospective interstripes move into nearby dark stripes or die in the interstripes generating a stripe/interstripe distinction [43]. During pattern development, new stripes and interstripes form sequentially, and seemingly as the result of the stripes and interstripes that appear before them. Starting from the onset of metamorphosis, at the Pelvic Bud stage (PB 7.2mm SSL) iridophores are visible in dense form along the horizontal myoseptum, marking the first X0 interstripe. Larval de-differentiated xanthophores re-differentiate in the X0 interstripe giving the X0 interstripe its yellow colour. From the X0 interstripe, iridophores proliferate and migrate to the edge of the X0 interstripe where they then transition from silver/dense to a blue/loose form. Meanwhile, metamorphic melanophores begin to appear either side of the initial interstripe, forming stripes 1D and 1V. By the Pelvic fin Ray stage (PR 8.6mm SSL), the X0 interstripe is complete and the adult primary melanophore stripes are partially formed: melanophores and loose iridophores occupy the 1D and 1V regions. From the stripes 1D and 1V, iridophores continue to migrate and proliferate to the edges of 1D and 1V where they next transition to dense form forming the beginning of interstripes X1D and X1V. By stage Squamation onset Posterior (SP 9.5mm SSL), stripes 1D and 1V are complete and interstripes X1D and X1V are partially formed: occupied by xanthophores and dense iridophores. This process of generating stripes and interstripes continues until all stripes and interstripes are formed at Juvenile+ stage (J+ 13mm SSL).

3.7.3 The initial iridophore interstripe provides a morphological landmark for pattern development

The *choker* (*cho*) mutant larvae, which lack a horizontal myoseptum [44], fail to show the initial focused production of iridophores marking the propsective X0 at the onset of metamorphosis [45, 44]. Instead, iridophores, which usually appear along the horizontal myoseptum at stage PB in WT fish, differentiate later in dispersed clusters above the myotomes throughout the body at stage PR. Following iridophore differentiation, xanthophores emerge in patches associated with dense iridophore regions across the domain, interspersed with melanophores [46]. As pattern development ensues, melanophore and xanthophore regions develop and separate, until a pattern emerges across the body by stage J+. This pattern consists of stripes and interstripes of the same spacing as normal WT fish, but which branches and twists in a labyrinthine manner [44]. The adult pattern indicates that whilst cues at the horizontal myoseptum are important for orienting the first interstripe, stripe and interstripe width is dependent on the interactions between melanophores, xanthophores and iridophores.

3.8. All three pigment cell-types are required to generate the adult pigment pattern

One key class of pigment pattern mutants are those missing a single chromatophore type: *ltk/shd* which lack iridophores, *csfr1a/pfe* which lack xanthophores and *mitfa/nac* which lacks melanophores. In all cases where a pigment cell-type is missing, normal adult pigment stripe formation does not occur, indicating that all three pigment cell-types are required to generate the pattern. Study of the single mutants in this class has

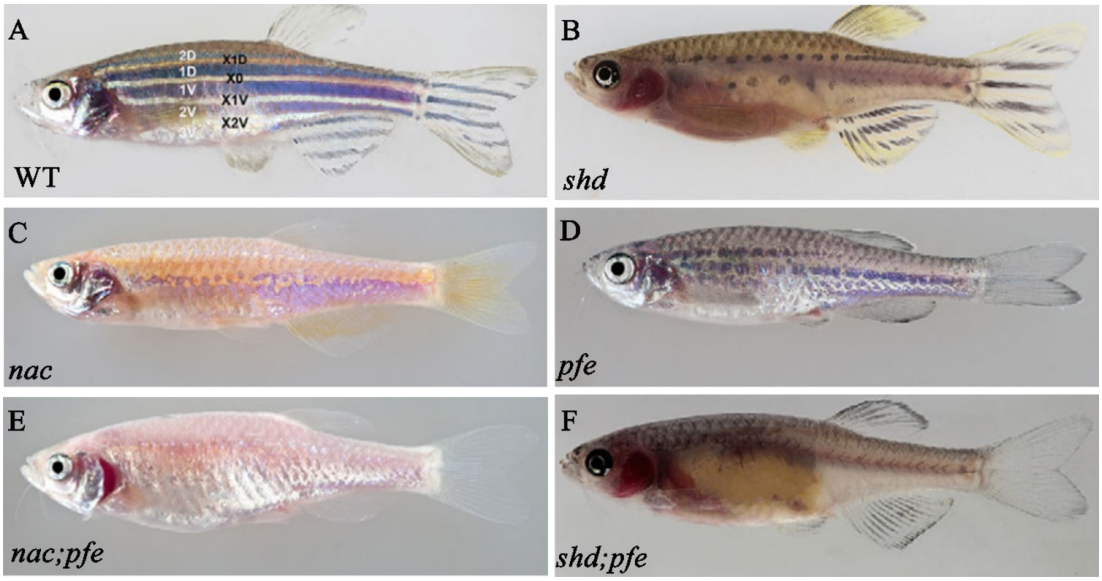


Figure 3-3: ***Mutations causing loss of one or more cell types.*** *(A) WT fish. (B) shd mutant fish lack iridophores and display spots of melanophores surrounded by xanthophores. (C) nac mutant fish lack melanophores and exhibit an extension of the X0 interstripe into blebs of xanthophores and dense iridophores ventrally, surrounded by loose blue iridophores. (D) pfe mutant fish lack xanthophores and exhibit melanophore spots as well as ectopic melanophores in interstripe regions. (E) nac;pfe mutants lacks melanophores and xanthophores and display a coherent sheets of dense iridophores (F) shd;pfe mutants lack xanthophores and iridophores and exhibit melanophores scattered uniformly over the fish. Images reproduced from (Frohnhofer et al., 2013 [46]) and licensed under CC-BY 3.0 (http://creativecommons.org/licenses/by/3.0); published by The Company of Biologists Ltd.*

been revealing regarding the roles of individual cell-types. A summary of the missing cell mutations is given in Table 3.2.

3.8.1 No iridophores (*ltk/shady*)

The strongest leukocyte receptor tyrosine kinase (*ltk*, also known as *shady/shd*) mutant alleles are early larval lethal, but escaper adults lack iridophores and display two stripes of melanophore spots surrounded by xanthophores [47]. In *shd* mutants, fewer metamorphic melanophores emerge during pattern metamorphosis and they remain more homogeneously dispersed over the entire flank. Later, following the widespread differentiation of xanthophores, melanophores aggregate into spots roughly in the positions of 1D and 1V (Fig. 3-3(B)). Since metamorphic melanophore number is reduced, this suggests that iridophores may play a role in promoting melanophore differentiation. Furthermore, aggregation of melanophores into spots in the absence of iridophores, suggests that iridophores are not necessary for enforcing strict stripe-interstripe boundaries. Inhibiting Ltk signaling during postembryonic stages results in the loss of adult iridophores, whereas ectopic iridophores are induced by Aklal2a overexpression and develop in moonstone, an *ltk* gain-of-function mutant [48, 49, 50].

3.8.2 No melanophores (*mitfa/nacre*)

The *melanocyte inducing transcription factor a* (*mitfa*, also known as *nacre/nac*) mutant lacks both larval and adult melanophores [51]. Adult *mitfa* mutants display an expanded X0 interstripe of dense iridophores and xanthophores with spots of xanthophores and dense iridophores more ventrally in a sea of loose iridophores [46] (Fig.

3-3(C)). Pattern development differs from metamorphosis. Instead of changing from dense to loose at the boundaries of the X0 interstripe, dense iridophores continue to spread further. Xanthophores differentiate associated with the dense iridophore regions. More ventrally, dense iridophores change to the loose blue form. Since in the absence of melanophores, xanthophores still localize in dense iridophore interstripe regions, this suggests that xanthophores may also be attracted to dense iridophores. As iridophores are able to transition between loose and dense form in the absence of melanophores, this suggests that xanthophores likely provide signaling cues for iridophore transitions independent of melanophores.

3.8.3 No xanthophores (*csf1a/c-fms/fms/panther/pfeffer/salz*)

In homozygous *colony stimulating factor 1 receptor 1a (csf1a*, also known as *pfeffer/pfe* and *salz/sal*) mutants, the development of xanthophores is strongly suppressed in larvae and abolished in adults [5, 52]. The adult homozygous *pfe* mutant displays spots of melanophores associated with dense iridophores most aligned with the normal stripe positions with loose iridophores and scattered ectopic melanophores in the interstripe regions (Fig. 3-3(D)). They also show an increase in S-iridophores and a 50% reduction in melanophore number. Iridophores appear in the X0 region as in WT, around stage PB although sometimes earlier, followed by the appearance of melanophores in stripe regions 1D and 1V in reduced numbers, suggesting that xanthophores are also important for melanophore differentiation. Iridophores proliferate and dense iridophores spread without transitioning into loose form, into 1D and 1V splitting the stripes into melanophore spots [46]. Removal of melanophores from the interstripes is mediated by xanthophores, as melanophores persist in the light stripe in *pfe* mutants. Since metamorphic melanophore number is reduced, albeit at a lower number than in the *shd* mutant, this suggests that xanthophores also play a role in promoting melanophore differentiation and/or survival. Furthermore, the appearance of melanophores in the interstripes, suggest that xanthophores are important for ensuring removal of melanophores from interstripe regions. As iridophores are able to transition between loose and dense forms in the absence of xanthophores, this suggests that melanophores may provide signaling cues for iridophore transitions independent of xanthophores. Meanwhile, fish heterozygous for *pfe* have a reduced number of xanthophores when compared with WT and show irregular and interrupted stripes [52]. Otherwise, stripes and interstripes are more clearly distinguished than those observed on fish homozygous for *pfe*, further suggesting that xanthophores likely play a role in maintaining strict stripe/interstripe boundaries.

3.8.4 Double mutants of *shd*, *nac* and *pfe*

Through combination of two of these missing cell mutants, mutants lacking two of the three chromatophore types have been generated [46]. In all of these double mutant cases, the remaining chromatophores homogeneously fill the domain, arguing against a patterning mechanism based on morphogens or physical prepatterning, and indicating that it is indeed the interactions between the three chromatophores that generate the pattern. Interestingly *nac;pfe*, which lacks melanophores and xanthophores, only display iridophores in dense form, whilst mutants lacking melanophores or xanthophores (*nac* and *pfe*, respectively) display iridophores in both dense and loose form. From this we can infer that both melanophores and xanthophores regulate the transitions of iridophores between loose and dense states.

Mutant	**Pigment cells present**	**Homozygous pigment pattern**	**Adult pattern**	**Deduced cell-cell interactions involved in pattern formation**
ltk/shd	X and M only.	Lack of S- and L- iridophores on the body of the fish. A reduction to 10-20% of metamorphic M in the adult fish	Spots of M aggregates, not necessarily in usual stripe region, surrounded by X.	M are more greatly repulsed by X than by M.
csf1a/pfe	M and I only	None or very few visible X. The area covered by M is strongly reduced	S-iridophores spread into stripes 1D and 1V and split the melanophore field into patches, which decrease in an antero-posterior gradient. M and I patches mix in the ventral half of the flank from stage SP onwards.	S-iridophores and M mutually repulse each other in the short range.
mitfa/nac	I and X only.	Complete lack of M. Increased number of I. [51] Number of X is variable.	An extended interstripe of X and dense S-iridophores with irregular borders in the region of X0, accompanied by spots of X and I ventrally separated by regions composed of a thin net of blue S-iridophores, as well as L-iridophores.	X promotes the localisation of I forming part of the aggregation process.
nac;shd	X only.	Lack of M and I.	X homogeneously fill most of the flanks in a dense layer	X can differentiate independently of the other cell-types.
shd;pfe	M only.	Lack of I and X	M homogeneously cover the flank. Density lowers with age	M can differentiate independently of the other cell-types, however, survival is likely dependent on X and M.
pfe;nac	I only.	Lack of X and M.	A dense layer of S-iridophores covers the trunk region, replaced towards anterior region by L-iridophores.	M and X mediate I shape transitions from dense to loose.

Table 3.2: *Mutations causing loss of pigment cell-types. M, X, I abbreviates melanophore, xanthophores and iridophores, respectively.*

3.8.5 Two missing pigment producing cells (*nac;pfe*, *shd;pfe*, *shd;nac*)

Double mutants missing two of the three pigment producing cell types, display the remaining cell type homogeneously across the body of the fish (Fig. 3-3(E)-(F)). A summary of all of the missing cell mutations and their effects are given in Table 3.2.

3.9. Regulation of cellular processes

3.9.1 Thyroid hormone

Thyroid hormone - best known for coordinating disparate cellular processes during amphibian metamorphosis as well as its functions in mammalian metabolism and neural development - has profound effects on all three major classes of adult pigment cells [36]. Zebrafish that lack thyroid hormone owing to transgenic ablation of the thyroid gland, or loss-of-function alleles of thyroid-stimulating hormone receptor, develop about twice as many adult melanophores as WT fish and lack visible xanthophores. No investigation has been made into the effects on iridophores. In these hypothyroid fish, melanophores proliferate inappropriately, whereas xanthophores are present but fail to acquire pigmentation. Adult hypothyroid fish observe broken stripes which become spots moving ventrally and dorsally away from the initial 1D and 1V stripes [36]. Conversely, a hyperthyroid mutant develops too many pigmented xanthophores and too few melanophores. Adult hypothyroid mutants display an expanded X0 interstripe with no stripes. Outside the initial interstripe are interspersed melanophores and loose iridophores, in a pattern similar to that observed in *D. albolineatus*. Recent findings suggest that the role of thyroid hormone is a key change between *D. albolineatus* and *D. rerio* [36]. Despite having opposite effects on melanophore and xanthophore population sizes, thyroid hormone promotes the maturation of both cell-types, as revealed by single-cell transcriptomics coupled with mutational and other analyses. For melanophores, thyroid hormone promotes a terminally differentiated, binucleated state in which proliferation ceases. For xanthophores, it drives acquisition of carotenoid-based pigmentation, making previously cryptic cells visible. Without thyroid hormone, melanophores remain young and continue to divide, whereas xanthophores stay hidden. Whilst, these results indicate the importance of the cellular process for pattern formation, more information is required about thyroid hormone effect on iridophores to discern pattern effects. For example, recent modelling studies indicate that the *D. albolineatus* phenotype, which has been identified as having evolutionary decoupling of the effects of thyroid hormone, is likely caused by changes to requirements for S-iridophore shape change during pattern development [28, 26].

3.10. Reduced iridophores in the adult

Another class of mutants have a partial reduction of iridophores. This consists of mutants affecting endothelin signaling, (*endrb1/rse, edn3a, edn3b*), as well as *mpv17/tra*. In the absence of endothelin signaling, iridophores differentiate near the horizontal myoseptum but fail to populate other regions of the hypodermis. Despite having differences in the underlying development - *tra* mutants develop iridophores which are subsequently lost, whilst *rse* mutants exhibit defective proliferation of iridophores - mutants of this class exhibit a similar adult phenotype to *ltk/shd*: reduction in the number of metamorphic melanophores and a reduced number of stripes, broken into spots. This suggests that iridophores play a role in pattern maintenance.

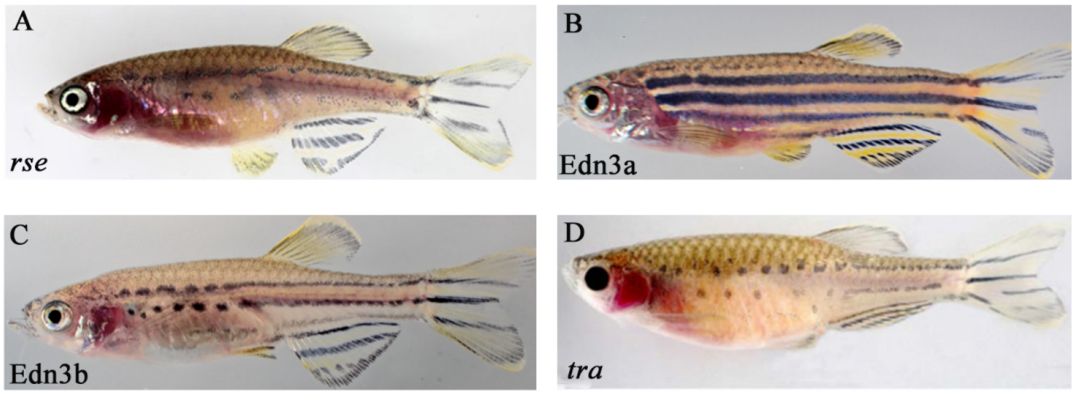


Figure 3-4: ***Mutations causing a reduction in the number of iridophores in the adult.*** *All fish (except for B) share a similarity with the phenotype of mutant shd which lacks both L- and S-iridophores. (A) rse fish lack all L-iridophores and some S-iridophores and exhibit two of five broken stripes at adulthood. (B) Edn3a mutant fish only lack L-iridophores and exhibit normal pattern development. (C) Edn3b mutant fish exhibit a dramatic decrease in the proliferation of S-iridophores and displays two broken stripes. (D) tra mutant fish exhibit an unusually high death of differentiated iridophores and displays two broken stripes at adulthood. Image (A) is reproduced from (Krauss et al., 2014 [53]), images (B)-(C) are reproduced from (Spiewak et al., 2018 [54]) and image (D) is reproduced from (Krauss, Astrinidis, Frohnhofer, Walderich, and Nsslein-Volhard, 2013 [46]). Images (A)-(D) are licensed under CC-BY 3.0 (http://creativecommons.org/licenses/by/3.0); published by The Company of Biologists Ltd.*

3.10.1 *alkal2a; alkal2b*

Ltk was an orphan receptor, but two candidate ligands for Ltk were identified recently. The orthologues of FAM150A and FAM150B (ALK and LTK Ligand 1 (ALKAL1) and ALKAL2 or augmentors, Aug) encode small secreted proteins [49, 50]. In zebrafish there are three augmentor orthologs, designated *alkal2a*, *alkal2b* and *alkal1*. Loss-of-function phenotypes in *alkal2a*; *alkal2b* double homozygous embryos are identical to those in *ltk/shady* mutants: the mutant embryos completely lack iridophores, suggesting that *alkal2a* and *alkal2b* are the primary Ltk ligands driving iridophore development in embryogenesis [50]. Prominent expression of all three *aug* genes in iridophores at 72 hpf equivalent to the expression of *ltk* itself [49], and of *alkal2a* and *alkal2b* in NCCs, e.g. near the eye, at 24 hpf [50], suggests that these ligands activate the receptor Ltk in a paracrine and/or autocrine fashion, although further characterization of the extent of expression is required. Embryonic overexpression of each, but especially Alkal2a, stimulates numerous ectopic iridophores in an Ltk-dependent manner, indicating that Alkal2a is perhaps more potent in stimulating Ltk activity in iridophores [49]. *alkal2a; alkal2b* double homozygous adult fish have severe iridophore reduction in the body like *ltk/shady* mutants do while a partial iridophore defect in the eye. In contrast, *alkal2b; alkal1* double mutants exhibit a severe reduction of eye iridophores with no significant body stripe defects, suggesting a larger contribution of alkal1 in the adult eye pigmentation.

3.10.2 *ednrb1/ednrb1a/ rse*

The *endothelin b1a receptor/rose* (*ednrb1a,* also known as *rose/rse*) mutants are defective for L-iridophores and some S-iridophores [55]. Furthermore there is a reduction to 30% of metamorphic melanophores. In strong mutant alleles of *rse* only the first stripes, 1D and 1V, develop, broken into spots. In weak alleles, stripes 2D and 2V

may also develop, with stripes broken into spots more ventrally (Fig. 3-4(A)). In adult zebrafish, *rse* is expressed by melanophore and iridophore lineages but becomes restricted to iridophores later. Moreover, cell transplantations have shown that *rse* acts cell-autonomously in iridophores, but not in melanophores, suggesting that iridophores supply an unknown factor required for melanophore development.

3.10.3 Edn3

Edn3a mutants have a normal pigment pattern, including stripes and iridophore interstripes but lack L-iridophores (Fig. 3-4(B)). Ednrb1a ligand endothelin 3b (Edn3b) mutants exhibit a dramatic decrease in the proliferation of S-iridophores in the skin (but not L-iridophores) whereas Edn3b overexpression results in iridophore population expansion and, indirectly, more melanophores (Fig. 3-4(C)). Edn3b has been shown to be specifically required for iridophore but not for melanophore development during early larval stages [53]. Edn3b mutants have S-iridophore and melanophore deficiencies but normal L-iridophores. The adult displays a similar phenotype to *shd*; two stripes of melanophores broken into spots, surrounded by xanthophores, except unlike *shd*, the stripes are more coherent and there are fewer breaks, especially towards the tailfins.

3.10.4 Mpv17/*tra*

The mitochondrial protein mitochondrial inner membrane protein (Mpv17, also known as *transparent/tra*) mutant phenotype is caused by a deletion in the Mpv17 [56] leading to the death of differentiated iridophores and a reduction in iridophore number throughout development. The adult *tra* mutant displays a phenotype similar to the *shd* mutant; two stripes of melanophores broken into spots in a sea of xanthophores (Fig. 3-4(D)). The reduction in iridophore pigmentation during pattern metamorphosis occurs gradually. The first X0 interstripe, forms normally, and unlike *rse* or *shd*, *tra* mutants initially develop two full stripes of melanophores (seen at 2 months). However, soon the silvery pigmentation is gradually lost and these stripes subsequently dissolve into spots by 6 months [56]. Cell transplantations have shown that *tra* acts cell-autonomously in iridophores, and the subsequent reduction in melanophores in the body occurs as a consequence of this loss. This is consistent with the idea that iridophores supply an unknown factor required for melanophore maintenance.

3.11. Genes affecting the pigment cell environment

In the pigment pattern field, the emphasis has been on how pigment cell interactions result in the emergence of the pigment pattern. However, another class of mutants have been shown likely to act non-cell-autonomously to the pigment cells. These mutations affect cell development indirectly by affecting aspects of the tissue environment during pattern formation. In particular, analysis of *bnc2* [57] and *karneol* mutants [53] point towards a crucial role of the tissue environment surrounding the pigment cells during stripe pattern formation. However, the apparent paradox of a cell-cell interaction system showing environmental influences is readily reconciled, since the two examples to date both control provision of environmental ligands acting through receptors expressed in the pigment cells.

3.11.1 *bnc2*

The gene basonuclin2 (*bnc2*) encodes a highly conserved, nuclear-localized zinc finger protein of unknown function that is required for normal stripe formation, but does not act in pigment cells themselves [41]. Strong mutant alleles of *basonuclin2* (*e.g.* $bnc2^{\text{bonaparte}}$) lack iridophores and almost all melanophores. They further display a few residual iridophores near the horizontal myoseptum, but a complete loss of stripes and interstripes on the body; in weak alleles (*e.g.* $bnc2^{poppy}$), the X0 interstripe is reduced and more tortuous, and 1D and 1V are broken up into asymmetric spots. During development, *bnc2* mutants exhibit extreme reductions in metamorphic melanophores and, unlike in WT fish acquire 'stripe' melanophores from persisting early larval lateral melanophores. This suggests that without local signals from differentiating metamorphic chromatophores, many more early larval melanophores can persist in the skin to adulthood. The *bnc2* gene regulates both *kitlga* and *csf1* expression, which are necessary for melanophore and xanthophore survival respectively, but it is difficult to be sure whether other targets are also important contributions to the phenotype [41].

3.11.2 Ece2/ *kar*

Endothelin-converting enzyme, (Ece2 also known as *karneol/kar*) mutants, have a phenotype similar to mutants with weak alleles of *rse* [53]. Adults display two stripes of spots of melanophores, with a reduction in iridophore and melanophore numbers. *kar* activates endothelin pro-ligands by proteolytic cleavage in the tissue environment, readily explaining the lack of cell autonomous requirement for *kar* activity in any of the chromatophores. *kar* promotes iridophore development in a non-cell autonomous manner [53] via a paracrine mode of endothelin signalling during pigment patterning facilitating the production of active Edns. Iridophores express the receptor *rse*, which in *kar* mutant embryos, cannot be fully activated due to the lack of processed ligand. Since the *kar* mutation affects iridophore production and melanophore numbers are also depleted, these results provide further evidence that iridophores supply an unknown factor required for melanophore development and maintenance.

3.11.3 Somatolactin

Somatolactin (SL) is a pituitary hormone of the growth hormone/prolactin family that to date has been found only in fish. SL function in pigment cell development was first suggested through study of the medaka colour interfere (*ci*) mutant, which was shown to encode SLa gene [58]. The *ci* adult fish show pale grey skin colour as a result of increase in leucophores in number and size and concomitant decrease in xanthophores while the *ci* larvae have a normal number of those pigment cells despite their abnormally-pale body colouration. These phenotypes indicate that SL predominantly acts after metamorphosis to promote xanthophore differentiation and to repress leucophore proliferation in medaka [59, 58]. Consistently, overexpression of SL results in more yellowish body colouration due to increased xanthophores and decreased leucophores. Receptors for SL (SLRs) have not yet been identified, except that salmon SLR has been identified as being highly similar to growth hormone receptors but biochemically specific to SL rather than to growth hormone. The contribution of this SLR to skin pigmentation has yet to be investigated. Likewise, it is not clear whether SL is likely to have a pigmentation role in zebrafish.

3.12. Genes required for early metamorphic melanophores

Early studies of metamorphic melanophore development distinguished two subsets of these cells, according to the time of their appearance: early appearing, which differentiate during the first week of metamorphosis, and late-appearing, which differentiate in the latter stages of metamorphosis. Single and double-mutant analyses indicate that early-appearing metamorphic melanophores are dependent on *kit/sparse* (*spa*) and late- appearing metamorphic melanophores, are dependent on *rse* and *pfe*. As expected, both early and late melanophores are absent when fish are doubly mutant for *spa* and either *rse* or *pfe*, whilst *pfe;rse* doubles show no significant further reduction in melanophore number beyond that seen in each of the single mutants [34]. However, as new evidence brings attention to the importance of iridophores and xanthophores in the generation of metamorphic melanophores, it becomes increasingly uncertain that late-appearing melanophores are directly dependent on *edrnb1* and *csf1ra*. The essential functions of both *ednrb1* and *csf1ra* are nonautonomous to melanophores, and implied mechanisms through which they act are only vaguely understood. We propose that the early and late populations are more likely distinguished by the indirect effect of a reduction of iridophores (*rse*) or the complete lack of xanthophores (*pfe*). Indeed recent mathematical modeling studies show that these reduced melanophore phenotypes in the cases of *rse* and *pfe* can be explained by simply reducing the number of iridophores and a complete lack of xanthophores respectively [28, 26]. Hence, we speculate that whilst early, *kit*-requiring metamorphic melanophores can be distinguished from late-appearing ones, the latter are those cells whose differentiation is stimulated by xanthophores and iridophores.

3.12.1 *kit/spa* and *kitla/slk*

The adult loss of function receptor *tyrosine kinase* (*kit*, also known as *sparse/spa*) and *kit ligand a* (*kitlga*, also known as *sparse-like/slk*) mutants are almost completely devoid of melanophores at the beginning of metamorphosis. Nevertheless, 1D and 1V stripes made up of a reduced number of melanophores appear later either side of the X0 interstripe. Adult stripes appear faded and more tortuous than those of WT fish. Furthermore stripes contain only 15% of the normal number of melanophores on the body [55]. Differentiation, migration and survival of early-appearing adult melanophore precursors are promoted by paracrine factor Slk signaling through the receptor tyrosine kinase Kita. During metamorphosis, *slk* is expressed in the hypodermis, and its ectopic expression induces melanophore differentiation in extrahypodermal locations. These results indicate that timely and numerous melanophore differentiation is required for pattern formation, but ultimately a very minimal number of metamorphic melanophore differentiation is required for some semblance of stripe development.

3.13. Genes required for metamorphic chromatophore development

Mutants with reductions in melanophores in the adult pattern but not during embryonic pattern development are important in identifying factors required uniquely to establish, maintain or recruit the latent precursors to metamorphic melanophores. Zebrafish *tuba813/puma* and *erbb3b/picasso* mutants both show reductions in the number of metamorphic melanophores, with only partial compensation through retention of early larval melanophores, and display disrupted striping in adulthood. Both mutants display

reductions in other chromatophores too, and it is currently unclear to what extent the metamorphic melanophore reduction results from cell-autonomous function of these genes.

3.13.1 *tuba8l3/puma*

The *tubulin, alpha 8 like 3* (*tuba8l3*, also known as *puma*) mutant exhibits substantial reductions in the numbers of metamorphic melanophores and xanthophores and an expanded X0 interstripe region with disrupted stripes. Whilst the early larval pattern is unchanged compared to WT, the *puma* mutant exhibits 73% as many xanthophores and only 31% as many melanophores compared to WT [27]. Tubulins are key components of the microtubule cytoskeleton, but the exact role of this variant is still unclear. The *puma* gene has been shown to play a role in promoting the development of late-appearing *pfe*- and *rse*-dependent metamorphic melanophores as well as the numbers and organization of early-appearing *kit*-dependent metamorphic melanophores [27]. Analysis suggests that *puma* acts cell autonomously to promote the expansion of pigment cell lineages mostly during metamorphosis but possibly also at even earlier embryonic stages, to neural crest-derived cells and possibly other lineages as well [27]. Pattern development in *puma* differs from the start of metamorphosis. Early lateral larval melanophores, which in WTs would usually die in response to local xanthophore signals [43], survive more frequently than WT early larval melanophores. Metamorphic melanophore births are negligible during early metamorphosis with a gradual increase towards the end of pattern metamorphosis 3 weeks later [27]. Metamorphic melanophores in *puma* differentiate close to the early larval melanophores and rarely migrate. More research is required to fully understand the effects of this mutation and how it aids our understanding with regards to pattern development.

3.13.2 Erbb3b/*picasso*

The erb-b2 receptor tyrosine kinase 3b (Erbb3b, also known as *picasso*) mutant displays two undulating stripes with breaks in the mid-trunk. *picasso* encodes Erbb3, and the phenotype identifies a set of Erbb-dependent progenitors of adult melanophores that are established in association with the peripheral nervous system. Partial loss of DRG cells through ablation of ErbB signaling by either pharmacological inhibitors or *picasso* knockdown results in defective formation of adult melanophores while formation of the embryonic/larval melanophores appears normal, suggesting that ErbB signaling is required at a very early embryonic stage for much later development of metamorphic melanophores. Pattern development in *picasso* is similar to *puma*: The *picasso* mutant develops very few metamorphic melanophores, particularly in the mid trunk and instead many embryonic/early larval melanophores persist into the adult during pattern metamorphosis.

3.14. Genes relating to cell shape

Mutations which affect the cell shape need not impact pattern development as such but may change the appearance of the stripes. However, the one example that has been characterized to date changes both the appearance of the stripes and the distribution of melanophores in the skin.

3.14.1 *bace2/wanderlust*

The adult *beta-secretase 2* (*bace2*, also known as *wanderlust*) mutant exhibits poorly defined stripes across the body tail and fins. Pattern development has not been described. These mutants have melanophores that are hyperdendritic and found ectopically in interstripes and other locations [60]. *bace2* is expressed by melanophores and encodes a sheddase that cleaves insulin receptors from the plasma membrane, thereby curtailing insulin/PI3K/mTOR signaling. The salient insulin gene, insulin b, is expressed in the head at EL stages and subsequently in muscle and other tissues, suggesting the likelihood of endocrine roles early and possibly endocrine and paracrine roles later.

3.15. Cell communication during pattern development

As seen in the previous sections, all three pigment cell-types are required to generate the striped pattern. Increasing evidence suggests that this is due to the fact the pigment producing cells rely on each other for signals relating to differentiation, death and cell movement. As a result, many cell, ablation and genetic studies have been conducted into cell communication during pattern development. Cell communication between zebrafish pigment cells are split into two types: Short range interactions which are defined as interactions between neighbouring cells that is likely regulated by filopodia or dendrites, and long range interactions which span around half a stripe width distance and are regulated by filopodia or airinemes.

3.15.1 Cell-studies into cell communication

Studies of cell-cell intractions *in vivo* and *in vitro* indicate that whilst there appears to be little to no homotypic cell-cell interaction, (note: this has not been assessed in iridophores), heterotypic interaction of xanthophores and iridophores with melanophores is likely to influence the movement and differentiation of metamorphic melanophores.

Melanophore-melanophore interactions

Despite the widespread aggregation of melanophores during pattern development, all cell studies indicate that melanophores do not interact in the short range. *In vitro* studies of extracted fin melanophores suggest that there are no obvious cell-cell interaction responses between melanophores in the short range [42]. *In vivo* ablation studies show that melanophores do not aggregate together [43] but instead may repulse each other in the short range.

Xanthophore-xanthophore interactions

There is also no evidence that xanthophores interact in the short range. Both *in vitro* and detailed electron microscopy analysis *in vivo* suggest that there are no obvious interaction responses between xanthophores in the short range [43, 42].

Xanthophore-melanophore interactions

In contrast, many cell-cell studies suggest that melanophores and xanthophores interact in the short range facilitated by filopodia. Firstly, *in vivo* studies into melanophore-xanthophore interactions during pattern metamorphosis indicated that when xanthophores appear in the interstripe regions at the onset of metamorphosis, melanophores initially

distributed randomly across the flank, either migrated away to the black stripe regions (3/4) or died (1/4) [18, 43]. In contrast, when xanthophores nearby isolated melanophores in interstripes were ablated, melanophores moved randomly, were less likely to leave the interstripe region and did not die, suggesting that melanophores are repelled by xanthophores in the short range and that xanthophores can promote the death of melanophores in the short range. This is consistent with findings of *in vitro* analysis of melanophores and xanthophores extracted from the fins. They observed a run-and-chase behaviour in the short range between melanophores and xanthophores whereby xanthophore contact-dependent depolarization of melanophores caused xanthophores to chase melanophores (surprisingly, in a specifically anticlockwise direction) as melanophores attempt to move away from the offending xanthophore. Since the melanophores are approximately four times faster than xanthophores *in vitro* (and faster in general *in vivo* [39], this gave rise to the idea that slow xanthophores chase fast melanophores [42]. Melanophores have also been shown to influence the shape of xanthophores. For example, it has been observed that xanthophores reorganise their protrusions, and become stellate when they encounter melanophores, even in the absence of iridophores [61, 38]. In some cases, xanthophores will locally rearrange to accommodate their increasing numbers [38].

Xanthoblast-melanophore interactions

During pattern formation, unpigmented xanthophore precursors differentiate in the prospective interstripe region, but also occur at lower densities in developing and completed stripes where they remain unpigmented or lightly pigmented as incompletely differentiated xanthoblasts [40]. Experiments by Eom *et al*, have shown that these xanthoblasts exhibit long (up to 5-6 cell diameters), thin projections with distinctive, membraneous vesicles at their tips, that extend and retract quickly (individual cells could extend several projections within 18 hours) [62]. These fast projections, named airinemes are unique to the xanthophore lineage, fast projections were rare for melanophores and not observed for iridophores. More remarkably, the airinemes associated with xanthophores appear to be regulated by macrophages [63]. Macrophages recognise airinemes by phospholipid phosphatidyl-serine (PS), an eat me signal that is commonly found on the outer leaflet of the plasma membrane on apoptotic cells. The current working model is that xanthoblast blebs present PS to macrophages, which drag melanophore-specific targeted airenemes out from the xanthoblast. Contact between the airinemes and the target cell activates Notch-delta signaling, which ultimately promotes the consolidation of melanophores into stripes. Without airineme production in the xanthophore lineage, melanophores persist ectopically in the interstripe instead of migrating into the stripe regions [62].

Iridophore-melanophore interactions

An increasing number of studies indicate that iridophores, in particular iridophores in dense form, also affect melanophores independently of xanthophores. During pattern formation, the center of melanophore cells on the melanophore layer are rarely seen directly above a dense-form iridophore, however, melanophores frequently settle adjacent to dense-form iridophores. Melanophores will migrate to or differentiate in dense iridophore-free sites, suggesting short range repulsion of melanophores by dense-form iridophores and that dense-form iridophores may inhibit the differentiation of melanophores in the short range via some unknown signal [41].

Iridophore-xanthophore interactions

Arrival of iridophores during metamorphosis leads to a compaction of the overlying xanthophores and in general xanthophore filopodia become shorter upon arrival of iridophores [38], suggesting that there may be short range attraction of xanthophores by iridophores.

3.15.2 Ablation studies to assess cell-cell communication

Ablation experiments, whereby a subset of cells are ablated on a developing or fully developed (and hence, strictly speaking, then addressing pattern maintenance rather than pattern formation) zebrafish and the subsequent differentiation of new cells is documented, can give insight into how cells, local to or in the long range, can promote or inhibit the differentiation of new cells. Ablation of a rectangular subsection of the adult zebrafish stripes and interstripes, including all visible melanophores, xanthophores and the majority of iridophores, leads to the subsequent generation of a labyrinthine pattern whereby stripes and interstripes of the same width stemming from the remaining stripes and interstripes curve in varying orientations. This regeneration indicates that stripe and interstripe widths are determined primarily by cell-cell interactions, but that orientation in the normal pattern formation process is driven by another, prepatterning process. More specific ablations restricted to individual cell-types can help us to understand specific cell-cell interactions during pattern development.

Melanophore-melanophore interactions

Ablation experiments, whereby in a fully developed stripe pattern, a subsection of interstripe is fully ablated as well as partial ablation of neighbouring stripes, suggest that melanophores may inhibit the differentiation of new melanophores in the long range [18]. When melanophores occupied neighbouring stripe regions there was less differentiation of melanophores in the ablated interstripe region than when the melanophores occupying the neighbouring stripe region at a half stripe width away were also ablated.

Melanophore-xanthophore interactions

Xanthophores promote the differentiation and survival of melanophores in the long range distance away [18]. When melanophores occupied stripe regions, fewer melanophores differentiated in the neighbouring interstripe region, than when melanophores occupying the stripe region were ablated.

Iridophore-xanthophore interactions

Ablation experiments indicate that iridophores promote the development of xanthophores. Ablation of interstripe iridophores prior to xanthophore development results in fewer xanthophores in regions from which iridophores were lost, although both iridophores and xanthophores can be recovered gradually during later development, indicating that early deficiencies can be compensated for [41].

3.15.3 Genetic studies into cell communication

Mutations which affect cell-signaling but not development of chromatophores are of particular interest, as they can provide insights into the molecular mechanisms of cell-cell interactions underlying stripe formation. Mutants of this form generate robust

patterns that deviate from the usual stripes. These mutations indicate the importance of cell communication during pattern development for stripe formation. Several mutants have been described in which dark stripes are broken into spots. These consist of mutants Cx 41.8/*leo*, Cx39.4/*luchs*, Tjp1a/*sbr* and Igsf11/*seu*. Other mutations have been described where the widths of the stripes are notably different to WT fish: Kir7.1/*obe*, *asterix/ase* and *Tspan3c/dali* mutants display fewer and broader stripes. In contrast, the *srm/ide* mutant shows notably thinner stripes.

Cx41.8/*leo*

The adult Connexin 41.8 (Cx41.8, also known as *leopard/leo*) mutant is chararacterised by a spotted pattern of melanophores and loose iridophores in a bed of dense iridophores and xanthophores on the trunk [6] (Fig. 3-5(A)). Pattern development in *leo* mutants begins to deviate from WT patterning from stage PR. Melanophores either side of the initial interstripe (which usually gather in the stripe regions in WTs), remain dispersed in *leo* mutants. By stage J, dense iridophores and associated xanthophores are dispersed in an irregular manner so that the X0 interstripe branches into the usual striped region breaking the melanophores stripes into patches. By stage J++ (16mm SSL) spots can be observed instead of stripes on the body of the fish. Cell autonomy studies demonstrate that expression of *leo* is required in melanophores and xanthophores but not iridophores [6, 64]. It has been suggested that *leo* forms heteromeric gap junctions, promoting cell-cell interactions between melanophores and xanthophores. Double-mutants for *leo* with missing-cell mutants (*shd, nac,* and *pfe*) reveal disturbed patterning in all cases, suggesting that the defective interactions of *leo* melanophores and xanthophores may subsequently influence the patterning of iridophores [6]. For example, the *leo;nac* double mutant which lacks melanophores displays dense iridophores covered by xanthophores and the absence of loose iridophore areas compared to *nac*. *leo;pfe* double mutants lack xanthophores and, instead of the spotted pattern observed in *pfe*, show melanophores randomly distributed across the body of the fish. *leo;shd* double mutants lack iridophores and also have an almost complete loss of melanophores on the body of the fish. It has been suggested that the *leo* mutant phenotype is largely caused by iridophores failing to undergo shape transitions from dense to loose form at the correct time and place [6]. *In vitro* studies have demonstrated defective cell-cell interactions between *leo* mutant melanophore and WT xanthophore cells. Unlike in WT cells, there is no melanophore repulsive response or xanthophore chasing movement observed, suggesting that *leo* is important for mediating these interactions. Interestingly, *in vivo* studies have shown that expression of *leo* in mutant *leo* melanophores (only) can partially recover striping, albeit thinner than normal stripes, [65], indicating that *leo* is particularly important for heterotypic interactions involved with melanophores.

Cx39.4/*luc*

The heterozygous Connexin 39.4 (Cx39.4, also known as *luchs/luc*) mutant zebrafish exhibit wavy stripes with breaks, whilst in homozygotes, stripes are completely dissolved into very small spots and individual melanophores (Fig. 3-5(B)). Pattern development in *luc* mutants begins to deviate from WT patterning from stage PR. Iridophores at the edges of the interstripes fail to change to the loose form and suppress melanophores. At stage J, dense iridophores form patches around aggregated melanophores. By stage J+, two undulating stripes either side of X0 can be observed with patches of dense iridophores more ventrally. *luc* is required in melanophores and xanthophores but not iridophores [6]. It has been suggested that *luc* also forms heteromeric gap junctions,

promoting interactions of melanophores and xanthophores. Mutant crosses of *luc* with missing cell mutants (*tra, nac*, and *pfe*) reveal disturbed patterning in all cases. For example, in *luc;nac* double mutants, which lack melanophores, both xanthophores and iridophores are more homogenously distributed than in *nac* alone, although the difference is subtle. In *luc;pfe* double mutant, which lacks xanthophores, melanophores are distributed more evenly and only rarely form small spots. In *luc;tra* double mutants, which lack iridophores, the spotted pattern of melanophore of *tra* is reduced, and only a few small melanophore spots are present, although this is a weaker phenotype when compared with *leo;tra*.

Igsf11/*utr15e1/seu*

The adult Immunoglobulin superfamily member 11 (Igsf11, also known as *seurat/seu*) mutant displays an irregular spotted pattern on the body of the fish and a spotted pattern on the tailfins [66] (Fig. 3-5(C)). Igsfl1 is required autonomously by the melanophore lineage and mutants exhibit defects in both melanophore survival and migration during pattern metamorphosis. Whilst melanophores in WT fish move at a rate of approximately 6μm/hour, Igsfl1 mutant melanophores move at a rate of 2μm/hour. Whilst the survival rate of melanophores in WT fish is >95%, the survival rate of Igsfl1 melanophores is around 70%. It has been hypothesised that *seu* could also play a role in mediating adhesion between differentiated melanophores. K562 human myeloid leukaemia cells transfected with either WT or mutant *seu* cDNAs exhibited different cellular behaviour *in vitro*. In particular, whilst cells expressing WT Igsf11 adhered to each other, forming large aggregates within two hours, cells expressing mutant Igsf11 did not. This hypothesis of melanophore aggregation during pattern development, however, has been disproved by time lapse imaging which shows that melanophores do not adhere to each other *in vivo* [64].

Tjp1a/*sbr*

The adult Tight Junction Protein 1a (ZO-1a/Tjp1a, also known as *schachbrett/sbr*) mutant, which exhibits delayed changes in iridophore shape and organisation caused by truncations in *sbr*, displays a checkerboard pattern of rectangular melanophore spots associated with loose iridophores, organised into stripes within a bed of loose iridophores and xanthophores (Fig. 3-5(D)). Pattern development in Tjp1a mutants deviate at stage PR. As new melanophores differentiate either side of the initial interstripe, iridophores, which usually transition from dense to loose at the interstripe boundary, remain dense for longer and spread over the melanophore stripe regions 1D and 1V. As a result the melanophore stripe breaks into a rectangular spotted pattern. Later in development, some dense iridophores transitions to loose form in regions densely populated with melanophores. Tjp1a is expressed in dense iridophores but down-regulated in the loose form. It is postulated that Tjp1a regulates the transition of dense iridophores to loose type. Though there is also a reduction in the number of melanophores, it has been shown that the phenotype is not caused by this reduction, but rather that the reduction is caused by the delayed transitions of dense iridophores to loose type [7]. Double mutant crosses with missing cell mutants demonstrate that functional Tjp1a is important for iridophore shape changes. For example, *sbr;shd* double mutants lack iridophores, but the pattern is indistinguishable from *shd*. However, *sbr;nac* and *sbr;pfe*, which lack melanophores and xanthophores respectively, display a greater spread of the dense iridophores across the domain. As a result, *sbr;nac* fish show

a reduced blue iridophore section on the flank compared to *nac*, and *sbr;pfe* displays smaller melanophore spots than in *pfe*.

kir 7.1/jag/obe

The adult inwardly rectifying potassium channel 7.1 (*kir7.1*, also known as *jaguar/jag* and formerly known as *obelix/obe* [67]. mutant has fewer (2-3 instead of 5) and broader stripes across the body tails and fins than WT *D. rerio*, and displays ambiguous stripe-interstripe boundaries [67] (Fig. 3-5(E)). Pattern development in *kir7.1* mutants deviate from WT development from stage PB onwards. Melanophores which appear either side of the initial interstripe fail to aggregate into stripes and instead persist in more ventral regions as well as within the initial interstripe itself. Xanthophores, which initially appear in the interstripe, later also emerge between the scattered melanophores [67]. *kir7.1* is expressed in melanophores and xanthophores. Transplantation and *in vitro* studies suggest that *kir7.1* is only required in melanophores and not xanthophores [67]. Currently, it is not clear how the function of Kir7.1 can be linked to the *jag* phenotypes. However, since the role of the Kir protein family is to maintain membrane potential and sustain ionic composition [68, 69] it is likely that Kir7.1 functions in cell-cell signalling. In particular, it is thought to be responsible for managing cell-cell interactions which enforce the clear boundaries between melanophore and xanthophores during pattern development [67]. Related to this, it has been postulated that the *jag* gene functions to mediate melanophore cell-cell interactions. *In vitro* studies have shown that the directed movement of melanophores away from xanthophores is inhibited in *jag* [42]. The plasma membrane of melanophores derived from *jag* mutants showed no electrical polarization and failed to retreat from xanthophores, indicating that Kir7.1 may set the resting membrane potential of melanophores. The loss of this repulsive interaction in Kir7.1 mutants appears to contribute to a failure of cell sorting *in vivo* [42].

tspan3c/dali

Heterozygous *tetraspanin 3c* (*tspan3c*, also known as *dali*) mutants have a reduced number of interrupted block-like stripes of melanophores. The *dali* homozygous mutants show thinner stripes than normal, with melanophores intermingled abnormally with xanthophores. The development of the *dali* mutant has not been described in the literature. *tspan3c* is expressed in melanophores and xanthophores. It is not clear exactly how the function of tspan3c can be linked to the *dali* phenotype. The role of the tetraspanin family is to act as scaffolds to anchor proteins at cell and organelle membranes within tetraspanin-enriched microdomains and hence to contribute to cell adhesion, migration, proliferation and cell-cell fusion. *In vitro* studies have shown that the *dali* mutants have reduced melanophore motility and disrupted melanophore-xanthophore interactions, but *dali* is not required for melanophore depolarization [70]. More information is needed to understand the role of *dali* in pattern development.

srm/idefix

Homozygous *spermidine synthase* (*srm*, also known as *idefix/ide*) mutants have fewer stripes, and these are less regular, usually narrower and often interrupted; the X0 interstripe is also expanded [12] (Fig. 3-5(F)). The *ide* phenotype develops during adult metamorphosis. The mutant phenotype becomes clearly visible at stage SP, in which the first interstripe is already noticeably wider than in WT. The full phenotype

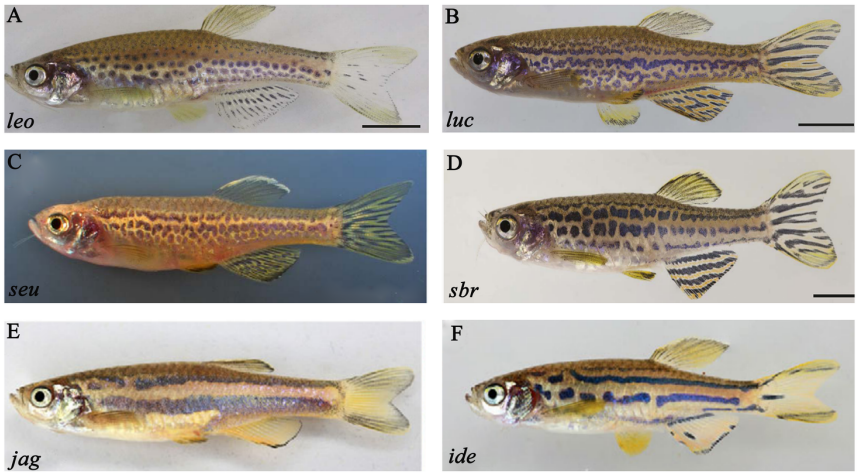


Figure 3-5: ***Mutations causing miscommunication between different cell types.*** *Fishes (A) leo and (B) luc are involved with the creation of gap junctions promoting cell-cell interactions between melanophores and xanthophores. (A) leo displays a spotted pattern, (B) luc display a undulating striped pattern. (C) seurat is required in melanophores and display an undulating striped pattern, (D) schachbrett is required for iridophore shape changes and displays a checkerboard pattern. (E) jaguar are involved in the creation of potassium channels and display wider stripes and interstripes. (F) idefix is involved in maintaining spermidine levels and display fewer less regular stripes. Images (A), (B) and (D) reproduced from (Fadeev et al., 2015 [7], Image (C) is reproduced from Eom et al [66], Images (E) and (F) are reproduced from Frohnhofer et al [12] and licensed under CC-BY 3.0 (http://creativecommons.org/licenses/by/3.0); published by The Company of Biologists Ltd.*

is developed by stage J++. The phenotype of the *ide* mutant is caused by a severe reduction in spermidine levels. *ide* is likely to influence the behaviour of the pigment cells by providing enough spermidine for the regulation of cell-cell communication [12].

3.15.4 Potential molecular mechanisms mediating pigment cell interactions

Analyses of the *leo* and *luc* mutants suggest that Cx41.8 and Cx39.4 interact to form heteromeric gap junctions and that these channels are responsible for some of the interactions that take place among melanophores, xanthophores and iridophores during pattern metamorphosis. Whilst *leo* and *luc* single mutants retain an organized but spotted pattern on the body of the fish, the body of the *leo;luc* double mutant is covered with dispersed melanophores associated with loose iridophores [6]. The rest of the flank around the melanophores is covered with dense iridophores and xanthophores. This suggests that in the absence of one connexin-partner, homomeric channels form, but retain only some functionality, whereas the absence of both connexin proteins in the double mutants results in a complete block to activity. Recently, Usui *et al* were able to identify the simplest gap-junction network required for stripe generation by generating mutants that were had defective *leo* or *luc* in either melanophores or xanthophores only and comparing the patterns that formed [71]. They discovered that Cx39.4 expression in melanophores was required but expression in xanthophores is not necessary for stripe patterning, whereas Cx41.8 expression in xanthophores is sufficient for the patterning, and Cx41.8 expression in melanophores might stabilize the stripes. However, more work is required to determine exactly how each of these connexins contribute. In *jag;leo* double mutant, there is a residual 1D melanophore stripe and X0 interstripe. Some melanophores are interspersed with xanthophores and iridophores

elsewhere [64]. For *jag*, previous analyses indicate that this gene is required in the melanophores to control their homotypic clustering, although it may also be involved in other cell-cell interactions [64]. The *jag;leo* double mutant suggests that at least one of the two cell-cell signaling pathways functions redundantly with a third signaling pathway to form coherent patterns. Tjp1a in iridophores may also be involved in cell communication with xanthophores, and/or melanophores. The *sbr* mutation enhances the phenotypes of both *luc* and *leo* mutants. The *leo;sbr* double mutant displays a decrease in the size of the spots compared to both *sbr* and *leo* [7]. The *luc;sbr* mutant exhibits an almost complete loss of melanophore clustering; the upper part of the body is covered with a layer of dense iridophores [7]. The very ventral region of the fish is covered in loose iridophores and some melanophores. Analysis of these double mutants suggest that Tjp1a and connexins do not act in a linear pathway to regulate pattern formation, but most likely work through different mechanisms. Absence or truncation of Tj1pa results in a delay between dense iridophores

1. receiving the signal from melanophores to switch to the loose form and
2. actually changing morphology.

The combination of this delay with the disrupted communication between melanophores, xanthophores and iridophores in *leo* and *luc* leads to an enhanced phenotype in both *leo;sbr* and *luc;sbr* double mutants. In double mutants where *ide* is combined with *leo, luc* or *jag*, a superimposition of both phenotypes is visible. That is, the characteristic spots (*leo, luc*) and widened stripes (*jag*) are combined with an extended X0 interstripe and slight reduction in melanophore stripe/spot width. Likely targets for spermidine are the gap junction components *leo, luc*, and *jag*. The phenotypic analysis of *ide* double mutants with *leo*, *luc* or *jag* suggests that *ide* does not exclusively affect one of these channels, and that it may affect all or even a different yet unknown channel. Interestingly, it has been shown that if Cx41.8 is expressed only in melanophores, but not in xanthophores, a phenotype very similar to *ide* is produced [72], suggesting that *ide* may play a function in heterotypic polarized gap junctions requiring Cx41.8.

3.16. Summary of mutations that cause defective patterns

In Table 3.3 we summarise the key features of mutants that have defective patterns from the last two sections alongside potential deductions about pattern formation. These comprise of mutations that cause fish to lose a cell type as well as mutations that cause damage to the communication between different cell types.

Gene name	**Protein type**	**Expression pattern in pigment cells**	**Mutant allele**	**Predicted allele type**	**Type defect**	**Adult stripe pattern phenotype (Body)**	**Pattern development deviations from WT**	**Deduced function of gene in pigment cell lineages**	**Deduced features important for pattern formation**
asterix	Unknown	Unknown	*ase*^{dtt269}	Unknown	Unknown	(Het) exhibit fewer and broader stripe (Hom) exhibit a stronger phenotype, with a further reduction in the number of stripes [5]	Not documented	Unknown	Not clear
baso nuclin-2	Zinc finger protein (Bnc2)	N/A - Hypodermal cells in contact with M, X and I	*poppy*^{j10e1}	Partial loss of function	Fewer metamorphic M, X and I Differentiated pigment cells die at high frequency	(Hom) Two partial stripes of M (arranged in very small spots)	Not documented	Promotes M and X development by regulating expression of the Kitlga and Csf1[41]	Not clear
			bona parte^{utr16e1}	Loss of function	Fewer metamorphic M, X and I, differentiated pigment cells die at high frequency	(Hom) Lack stripes on the body	Lack early appearing metamorphic I, but develop dispersed metamorphic M later [57]		

choker	*meox1*	N/A	*cho*^{tm26}	Loss of function	Loss of the horizontal myoseptum [45]	Stripes and interstripes are normal width but occupy a labyrinthine pattern	M appear evenly dispersed, I and X emerge in patches, but then aggregate into stripe-like arrangements, and separate from M regions [46]	N/A	I occupying the horizontal myoseptum play a role in aligning the stripes
dali	Transmembrane scaffolding protein Tetraspanin 3c [70]	M and X	*dali*^{wpr21e1}	Missense substitution that causes ER-retention and incomplete glycosylation of Tspan3c protein	Reduction in M movement, M-X interaction impaired	(Het) broken stripes with clear boundaries between M and X (Hom) fewer, broader stripes that contain M and X	Not documented	Cell-cell interactions between M and X,	M-X cell-cell interactions are important for pattern formation
duchamp	Unknown	Unknown	*duc hamp*^{utr19e1}	Unknown	Unknown	(Het) fingerprint-like spots across the body, 45% of M remain (Hom) exhibit more dispersed M than het [73]	Not documented	Unknown	Unclear
Edn3	Ligand, endothelin signalling	I	*edn3ab*^{1282}	Loss of function	Lack peritoneal iridophores	Normal stripes and interstripes, rosy cast to the ventrum [54]	Not documented	edn3 promotes iridophore development in *Danio* [54], required for larval iridophore development [53]	Unclear
			edn3bb^{1283}	Loss of function	Deficiency in hypodermal iridophores	Two stripes are broken into spots	Not documented		

idefix	Enzyme	spermidine synthase [12]	Not required in pigment cells	*ide*^{t26743}	Loss of function	Likely cell signalling interrupted	(Hom) The light stripe area is wider than in WT and only two dark stripes develop narrower than WT with frequent interruptions	Boundaries between the first interstripe and stripes are less regular from stage SP	Spermidine likely required for gap junctions and hence cell-cell signalling Unclear
jaguar	K+ channel *kcnj13*; (previously Kir7.1) mutation [67]	M and X lineage	*jag*^{b230}, *obe*^{tc271d}, and *obe*^{td15}	Loss of function	M-X interaction impaired, M constantly depolarized and is not repelled by X [42]	(Het) fewer and broader stripes on the fins and some incomplete stripes on the body (Hom) broader interstripes and thinner stripes, boundaries between M and X are blurred	M fail to aggregate into stripes and instead persist into more ventral regions as well as within the initial interstripe itself, X later emerge between the scattered melanophores	Cell-cell interactions between M - X	M-X cell-cell interactions are important for maintaining stripe/ interstripe boundaries
karneol	Enzyme Endothelin-converting enzyme 2	No chromatophores [53]	*kar*^{ltNO046}	Premature stop and the loss of the C-terminal peptidase domain containing the catalytic centre of the enzyme	Reduction of I and M [53]	Reductions in iridophores and melanophores [53]	At stages PB and PR, *kar* mutants show only slight reductions in I; from PR to SP reduction of I becomes more apparent, I do not extend properly into dorsal and ventral interstripe regions, occasionally, remnants or patches develop; from stage SP onwards the number of M is reduced compared to WT	*kar* promotes iridophore development in a non-cell autonomous manner	Unclear

kit	Receptor Kit signalling	eNCC and M	*spa*^{b5}	Loss of function	M precursors divide on average only ≈1.75 times between leaving the NC and establishing the early larval pattern	One-third reduction in M, with I and X less affected, stripes appear faded and more tortuous than those of WT fish, stripes contain only 15% of the normal number of M	For the first week of pattern metamorphosis M are almost completely absent, M only begin to differentiate during middle and late stages of metamorphosis	Promote survival, migration and differentiation of melanoblasts at all stages of development [74]	Timely M differentiation important for pattern formation
kit lga	Ligand Kit signalling	eNCC and M	*sparse like*^{tc244b}	Loss of function	Reduction of I, increase of M, late larval and metamorphic M are absent [34]	(Het) wave-like boundaries between stripes and interstripes, sometimes break into spots [34], (Hom) spotted pattern [34]	Almost completely devoid of M at the beginning of metamorphosis	Early NCC specification and suppression of *mitfa*	Metamorphic M are important for generating the stripes

leopard	Gap Junction Protein Connexin 41.8 [75]	M and X	leo^{tq270}	Missense mutation		(Het) Elongated spots instead of stripes/ some spots (Hom) Spot pattern	Not documented	Gap-junction formation between M and X [6, 75].	M-X interactions are important for maintaining M-X boundaries
			leo^{t1}	Nonsense mutation (Recessive)	M scarcely show repulsive response to X [34, 42]	(Het) No change (Hom) Spots	Not documented		
			leo^{tw28}	Missense mutation		(Het) Undulating stripes (Hom) Spot pattern.	M stay more dispersed instead of arranging in a stripe like pattern either side of the interstripe, dense I covered with X disperse in an irregular manner in the mutants to form a wider first X0 interstripe, branching into the stripe region and suppressing M dorsally and ventrally to form irregular patches		
			leo^{tk3}	Loss of function		(Het) Undulating and disrupted stripes (Hom) Individual M that hardly cluster together are still present, mostly associated with loose I			

luchs	Gap Junction Protein	Connexin, 39.4 [6] M and X, but not I	*luc*^{tXA9} and *luc*^{tXG1}	Loss of function	I appear almost entirely in dense form	(Het) undulating and disrupted stripes on their flanks (Hom) Individual M that hardly cluster together are still present, mostly associated with loose I	M ventral to the first interstripe are more scattered at stage PR. The second interstripe fails to form in the ventral region Later at stage J++ irregular stripes and spots become visible [6]	Cx39.4 form heteromeric gap junction channels in the plasma membranes of M and X which may facilitate communication with I	Suggest M and X interact with I to signal shape transitions
oberon	Unknown	Unknown	*obe*^{j198e1}	Not clear	Unknown	Pale stripes and interstripes	Not documented	Unknown	Not clear
rose	Receptor Endothelin receptor B (Ednrb1)	X precursors and I	*rse*^{tLF802}	Strong loss of function	Severe deficit of I and M [46, 55]	2 of 5 stripes develop. Fewer I and M than WT [53].	Appearance of X in *rse* mutants is delayed in X0, M differentiate near to X0 interstripe in reduced numbers, X fill the space between melanophore aggregates	*rse* may be responsible for the expansion of I	Dense S- iridophores may exert a long-range positive effect on M [46]. Lack of I, and to a lesser extent of X, is correlated with a strong reduction of M numbers
			rse^{tAN17X}	Weak (partial loss)	Deficit of I	2 of 5 stripes develop. Fewer I and M than WT [53]			
			rose^{b140}	Premature stop codon	Fewer I and M	2 of 5 stripes develop, 1D stripe forms normally, 1V stripe is broken into spots [55].			
			rose^{j3e1}	Missense substitution	Fewer I and M	4 stripes develop, 1D and 1V form normally, 2D and 2V are broken into spots [55]			

seurat	Cell surface protein that mediates adhesive interactions, Immunoglobulin superfamily member 11 (Igsf11) [66]	M	$seu^{utr15e1}$	Missense substitution (strong)	M migrate at 1/3 of normal speed, survival of M is around 70% (instead of 95%)	(Hom) Irregular spots of M	Not documented	Promote survival and migration of M lineage [66]	M survival is important for pattern formation
			$seu^{twp15e2}$ (weak)	Mutations at the boundary between the predicted signal sequence and the beginning of the first immunoglobulin domain	Reduction in number of M	Two stripes of M with one stripe broken into spots ventrally	Not documented		
schachbrett	Tight Junction Protein 1a [7]	Dense I	$sbr^{tnh009b}$, sbr^{twl1}, sbr^{twl4}, sbr^{twc2}	Partial loss of function	Delayed transition of I from loose to dense	Interupted undulating dark stripes of normal arrangement and width, M stripes are broken into rectangular spots	From the initial interstripe, I spread in a coherent sheet over M. Later some of the dense I change to loose type	Promotes transition of dense I to loose form	Cell shape transition of I from dense to loose are important for pattern formation

picasso	Receptor Epidermal growth factor receptor (EGFR)-like tyrosine kinase ErbB signalling [76]	eNCC and M	*pica sso*^{wpr2e2}	Loss of function	Persistence of early larval melanophores, deficient in metamorphic M particularly in the mid-trunk [77]	(Het) WT pattern development ensues [77] (Hom) Two (of 5) incomplete stripes form [77]	Early larval M persist and are incorporated into an incomplete stripe, additional metamorphic melanophores differentiate where persisting embryonic/early larval M contributed to the adult stripe	*erbb3b* promotes the development of latent precursors [77]	Metamorphic M are important for pattern development
puma	*tuba8l3*	eNCC, M	*puma*^{j115el}	Recessive loss of function	Persistence of early larval melanophores Reduction in M birth particularly during early metamorphosis with a gradual increase during terminal pigment pattern metamorphosis [78]	(Hom) Fewer stripes with irregular borders and many breaks, 73% of X and 31% of M [78]	Persistence of early larval M which move to form irregular striped patterns, during late metamorphosis, generation of metamorphic M into disorganised stripes	Early M precursor proliferation	Timely M differentiation important for pattern formation

Table 3.3: *A summary table of the all of the mutations that affect body stripe pigment pattern formation*

3.17. Dorso-ventral countershading

Alongside the striping pattern, a more general aspect of the pigment pattern common to almost all vertebrates, and widespread across the animal kingdom, is displayed in zebrafish. Thus, the dorsal region is noticeably darker and the ventrum paler, a phenomenon known as Dorsal Ventral Countershading (DVC). This evolutionary patterning mechanism is thought to be a basic camouflage mechanism from predators or prey positioned directly above or below the animal. From above, fish exhibiting DVC camouflage with the dark sea and from below, the bright sky. Countershading in fish is achieved by a patterned distribution of the pigment cells, with the light-absorbing and light-reflecting chromatophores mostly distributed in the dorsal and ventral areas, respectively [79, 80]. Experimental data in fish and amphibian species suggest that DVC is regulated by a combination of a melanisation simulating hormone (MSH) with a melanization inhibition factor (MIF) [81]. MSH is a hormone produced in both the ventral and dorsal regions that promotes the differentiation of melanophores and possibly inhibits iridophore proliferation. The identity of MIF has long been unknown, but MIF is thought to be locally produced by cells in the ventral skin and acts as an antagonist to MSH, inhibiting melanoblast differentiation and stimulating iridophore proliferation in this region [79, 81]. Recent studies have identified Agouti-signalling protein 1 (*asip1*) as the fish MIF [10, 8]. A priority for future research will be to understand how the balance of Asip1 and MSH results in the differential abundance of pigment cells in different body regions.

3.17.1 Scale development

Contributing to the darker dorsal surface is the association of pigment cells with dorsal scales. Development of scales begins posteriorly over the lateral surfaces of the tail at 8.1 mm SL and proceeds from this location towards the anterior [82]. By juvenile stages around 11mm SL, scales are fully formed. Pigment cells generated from adult pigment stem cells, and which likely travel along the nerves lining the scales [83], form two rows along each scale edge [31]. By adulthood multiple rows of pigment cells line the distal edges of the scales. The distribution of the pigment cells correlate with the spatial position dorso-ventrally on the fish. Thus, scales located dorsally contain all of melanophores, xanthophores and iridophores. As a result the dorsum appears dark. In contrast, the ventral scales do not contain any pigment cells [8]. Since melanophores are also virtually absent in the belly skin, the ventral region appears white as a result of the high number of iridophores in the abdominal wall.

3.17.2 Genes relating to dorso-ventral countershading (DVC) Agouti signaling

The melanocortin system consists of four endogenous antagonists of MSH signaling, which together form the agouti signaling protein family; *agouti signaling protein 1* (*asip1*), *agouti signaling protein 2* (*asip2*), *agouti related protein 1* (*agrp1*) and *agouti related protein 2* (*agrp2*). Of these, the most well understood Agouti signaling molecule in zebrafish is *asip1*. *asip1* is a secreted protein that serves as an important component of dorsal ventral countershading by acting as an endogenous antagonist for MC1R [9, 84, 85]. Expression of *asip1* represses melanization and mitfa expression ventrally, reducing the number of melanophores cells, although analyses of *asip1* mutants indicate it may also have roles in inhibiting the production of other cells too. The zebrafish

mutant *asip1*, which likely results in a partial loss of *asip1* function, displays a wider 2V stripe and extra 3V stripe ventrally. The belly, branchiostegal, jaw, ventral skin and operculum regions are hyperpigmented with melanophores and xanthophores and hence appear darker than in WT fish [8]. The internal abdominal wall, which usually contains a high number of iridophores, has a reduced number and appears more yellow (due to increased xanthophores) than WT. Pigment cells on the scales are also affected. At 210 dpf, ventral scales in *asip1* mutants show ectopic melanophores, xanthophores and iridophores; in contrast, dorsal scales are unaffected. In contrast, transgenic fish ubiquitously overexpressing *asip1* have an almost normal striped pattern, although stripe 2D is thinner. However, the DVC is reduced, with fish showing a hypopigmented dorsum, and retaining a light belly. In 210 dpf *asip1* transgenics dorsal scales are hypopigmented compared to WT and *asip1* KO. This is consistent with *asip1* overexpression phenotypes in other teleost fish; e.g. in turbot and sole, injection of *asip1* caused localized lightening of pigment [10]. Analyses of these mutants indicate that aside from inhibiting melanophore differentiation, *asip1* may also play a role in inhibiting the production of other pigment cells ventrally.

3.17.3 Melanocortin receptors

The teleost fish melanocortin system consists of five melanocortin receptor subtypes (MC1R, MC2R, MC3R, MC4R, MC5R) [9]. Zebrafish has six, because Mc5r is duplicated, as MC5Ra and MC5Rb [86]. The melanocortin 1 receptor (MC1R) is the central melanocortin receptor involved in vertebrate pigmentation. The zebrafish loss of function mutant of MC1R displays a similar phenotype to the *asip1* mutant. It is characterized by a darker ventrum than WT. The branchiostegal, jaw, ventral skin and operculum regions are hyperpigmented with increased numbers of melanophores and xanthophores. The abdominal wall appears overall more yellowish than WT [87]. Whilst most of the striped pigment pattern is not changed compared to WT, MC1R KO fish develop a thicker 2V stripe and an additional 3V stripe in the ventrum compared to WT. The scales are also affected in MC1R KO. The dorsal scales exhibit a strong reduction in the number of melanophores. The ventral scales which usually lack pigment cells, exhibit both melanophores and xanthophores. The MC1R KO still retains pigmentation ventrally, revealing that *asip1* interaction is more complex than simply blocking of the constitutive activity of MC1R. Rather, this data suggests that melanophore lineage specification and differentiation in the dorsal skin are not dependent solely on MC1R activity, and thus, some other receptors, presumably other melanocortin receptors, interact with *asip1*. MC2R, MC3R, MC4R, MC5Ra, and MC5Rb bind many MSH ligands with varying affinity [88].

3.17.4 Suggested molecular mechanisms

Asip1 is an endogenous antagonist and inverse agonist of melanocortin 1 receptor (MC1R). Many of the effects of dorsal-ventral countershading are likely mediated through MC1R, via ligands *asip1* and *agrp2* [9, 10]. Different theories exist regarding the result of *asip1* antagonism of the MC1R pathway in fish. Thus, it may 1) promote melanin synthesis and/or 2) switch cell fate from iridophores to melanophores, with further work needed to assess these definitively. Future work is required to assess the roles of the other MCRs in this process.

3.18. Conclusions

While there has been enormous progress in our understanding of pigment pattern formation in the zebrafish, and recent modelling studies have indicated that the biological mechanisms identified are nearly sufficient to explain pigment pattern formation up to juvenile stages at least [28, 26] there do remain some key aspects deserving of further experiemtnal investigation, in particular the likely role for L-iridophores in pattern maintenance. Furthermore, further investigation of the mutants defined here, to directly quantitate the interactions between cell-types and underlying molecular mechanisms that they define, promises a remarkable understanding of this fascinating process. Understanding of the evolution of these pigment patterns is the next major challenge.

3.19. Chapter Conclusions

This chapter provides an overview of the current known biology of zebrafish stripe formation. These findings will have a major role in determining the rules of our mathematical model of zebrafish stripe formation given in Chapter 4.

This chapter illustrates that, whilst most of the process of pattern formation is well described, many aspects are still not well understood. For example, whilst there are extensive studies on the interactions of the cell-types melanophore and xanthophore, there is limited available information about iridophores: despite being identified as being crucial for generating a striped pattern. The precise functionality of many genes involved in pattern development, particularly those involved in cell-communication such as *leo* and *luc* are also still unknown.

One way to address such questions is to build mathematical models of the pattern development and use these models to generate and/or test hypotheses. In the next chapter we will describe all previous mathematical models of zebrafish stripe formation as well as their contributions made to our understanding of the process. Later, in Chapter 4, we show how we can use our own mathematical model to address the remaining questions.

Chapter 4

A literature review of the mathematical modelling of zebrafish pattern formation

4.1. Overview

In the previous chapter we reviewed the results of all experimental investigation into the zebrafish stripe formation. In this chapter we provide a literature review of all mathematical models of zebrafish pigment pattern formation.

The aim of this chapter is to motivate the use of mathematical modelling in the context of zebrafish stripe formation as well as giving an overview of the different approaches that have been undertaken previously. In doing so, we will both outline the successes and major findings of these models, as well as their limitations. These findings will motivate the generation of the comprehensive biology-driven model of zebrafish stripe formation that we will build in Chapter 6.

We will start by describing the motivation for constucting a mathematical modelling study of zebrafish stripe formation. We distinguish two types of modelling that have been conducted in relation to this topic: component-level models, which only consider components of an overarching system, and comprehensive models, which aim to consider the system in its entirety. Of the two categories, we will break comprehensive models down further into continuum models, which model cells as continuously valued densities and discrete models, which represent cells as individual entities. At the end of the chapter, we provide a summary of all the mathematical models along with their major insights and contributions to the field of pigment pattern formation more generally.

In the context of the thesis, the role of this chapter is to provide the relevant background for Chapter 6, wherein we build a novel comprehensive discrete mathematical model. By outlining previous modelling efforts we are able to distinguish novel ideas that have resulted from using our model.

4.2. Outline

This chapter is made up of five sections. In Section 4.3 we will outline the motivation for generating mathematical models in order to understand zebrafish pigment pattern formation. In Sections 4.4 and 4.5 we will outline the major findings from all component-level and comprehensive mathematical models, respectively, of zebrafish pattern formation from the literature. We conclude in Section 4.6 with a summary of all mathematical models aforementioned.

4.3. Introduction

Given the large number of studies into the biology of zebrafish pattern formation over the past decade, it has become increasingly difficult to distinguish between the biological processes that are essential for correct pigment pattern formation from those that are merely interesting observations of the patterning process. One way to make this distinction is to build detailed mathematical models in which each component can be interrogated in order to establish its importance to the pattern generating process. As a result, there have been an increasing number of mathematically underpinned studies of zebrafish pigment pattern formation. Such mathematical approaches are beneficial to our understanding since they allow for the testing of biological hypotheses that may be unreachable experimentally giving valuable theoretical insight and potentially suggesting novel lines of experimental investigation.

Models for pigment pattern formation can typically be classified into one of three types; component-level models; comprehensive continuous models; and comprehensive discrete models. By component-level models we refer to models in which only a very minor subset of patterning-related behaviours are investigated, with no attempt to reproduce the organism-level pattern, these might look for example, specifically at features such as cell movement or differentiation. This is in contrast to comprehensive models which attempt to represent the patterning process itself, albeit with many important biological details ignored. These three model types are summarised in sections 4.4, 4.5.1 and 4.5.2. An overview of all the models discussed in this review chapter is given in Table 4.2.

4.4. Component-level models

Component-level models only consider *components* of the overarching system. So far, there have been four component level models considered for pigment pattern formation. Most of these studies focus on individual interactions between melanophores and xanthophores. These models are important for demonstrating key features that may be applicable to many types of pigment pattern formation, but are not comprehensive enough to describe the full system. The major findings of these models in relation to zebrafish stripe formation are that:

1. **Heterotypic repulsion and homotypic attraction between melanophores and xanthophore could generate stripes.** In 2008, Caicedo-Carvajal *et al* [23] undertook an *in silico* mathematical investigation into homotypic and heterotypic attraction and repulsion combinations between two species that could generate similar stripes to those seen in wild-type zebrafish.

2. **Community effects may be a viable patterning mechanism.** In 2011, Bloomfield *et al* [17], investigated the idea of a 'community effect'; the effect observed when stem cells acquired a greater affinity to differentiate into the cell type comprising the greatest proportion of cells in the local vicinity than the other cell type. (Note that this is different to 'the community effect' as defined by Gurdon in 1988 [89]).

3. **A chase and run mechanism (as observed between cells extracted from zebrafish) between two cell types is not sufficient for pattern formation.** In 2014, *in vitro* findings by Yamanaka *et al* [42] showed that melanophores and xanthophores extracted from the fins move according to a chase-run mechanism,

in which xanthophores chase melanophores. Inspired by these findings, Woolley *et al* [13] and later Painter *et al* [24] also consider this mechanism for pattern formation.

In the next subsection we give a more detailed description of each of these three component level modelling efforts.

4.4.1 Heterotypic repulsion and homotypic attraction between melanophores and xanthophore can generate stripes

In 2008, Caicedo-Carvajal *et al* [23], inspired by findings that had elucidated the relation between cellular interactions and surface patterns in identified zebrafish mutants [78, 64], performed a thorough mathematical investigation into heterotypic and homotypic attraction and repulsion between two cell species. They used a stochastic particle-based model developed by Shinbrot in 2006 [90].

In this model, cells are represented as disks (or spheres in three dimensions) that move randomly throughout the domain (*i.e.* according to a random walk) and interact with each other via attraction or repulsion forces described as follows.

1. Homotypic or heterotypic cell-cell attraction. When the distance between the surfaces of two cells that exhibit attractive forces is less than or equal to a predefined maximal distance they attract each other with a force proportional to the distance between cells. This attraction mechanism simulates the membrane tension of two cells connected by membrane-bound ligands.

2. Homotypic or heterotypic cell-cell repulsion. Cells exhibiting repulsive forces, repel each other when cells come into contact, moving in opposite directions upon collision.

Using this simple framework, the authors show how varying the interactions and their strengths theoretically leads to a range of patterns, including those similar to those seen in *Danio rerio* and mutants of *Danio rerio*. For example, they show that spots similar to those exhibited in the *Danio rerio* mutant *leo* can be generated when there is heterotypic repulsion of the same magnitude between opposing cell types and homotypic repulsion that is slightly greater in one species than the other. Alternating stripes such as those seen in *Danio rerio* form when there is homotypic repulsion and heterotypic attraction of the same magnitude. In particular, alternating stripes of like-cells separated by spaces without pigment cells occurs when heterotypic repulsion is greater than homotypic repulsion, although in this case, striping was more similar to heterozygous *Danio rerio* mutant *dali*, as stripes did not exhibit consistant alignment in one direction.

Whilst many of these theoretical pattern formation mechanisms have been shown not to occur in zebrafish (for example, it has since been shown that the mutation in *leopard* does not cause mutual repulsion in cells [42]), the generality of the model makes it applicable for understanding of other patterns. Therefore, while a lack of biological justification limits the applicability of this study for understanding pattern formation in zebrafish, as a purely theoretical study it provides many potential hypotheses and avenues of investigations for biological experimentation for pattern formation in other animals.

4.4.2 Community effect may be a viable patterning mechanism

Community effect is defined as the preferred differentiation and increased survival rate of cells in a region which has a high proportion of the same cell type. During zebrafish stripe formation, melanophores die in areas of high xanthophore density due to a community effect [43].

In 2011, Bloomfield *et al* [17] investigated this phenomena in a theoretical study using two species integro-differential equations. They modelled two species; A and B that each had a constant death rate and a differentiation rate determined by a function f whose argument is the proportion of the cell type in the local vicinity (found by integrating over space) and outputs the according differentiation rate. Therefore, in the case of a community effect, f should be monotonically increasing - as the higher the proportion of cells of the same type, the greater the rate of differentiation. They consider the following three models:

1. **The control model (cell autonomous model):** In this scenario, differentiation into the two different cell types occurs at a constant rate independent of the local environment. That is, f is a constant function.

2. **The single cell model**: f is linear and increasing, this is a weak community effect model.

3. **The community effect model:** f is a non-linear continuous step-like function.

Upon performing linear stability analysis in one and two dimensions they showed that without dependence on the local environment (as in the cell autonomous model), pattern formation does not occur.

Similarly the weak community model could not produce permanent patterns. The only model to generate striped patterns was the strong community effect model, which produced stripes, or spots provided the death rates of the two cell types was equal (otherwise one cell type dominates). Again, this model provides some insight into how a community effect system could provide a pattern formation mechanism, however as a component-level model it lacks biological insight.

4.4.3 Chase and run between melanophores and xanthophores is not sufficient for pattern formation

In a 2014 review by Woolley *et al* [13] of the *in vitro* analysis provided by Kondo [42], the authors generate a simplistic stochastic two dimensional run and tumble model of melanophores and xanthophores. At each time step, xanthophores and melanophore move randomly with some bias such that xanthophores chase melanophores, that is melanophores move away from nearby xanthophores and xanthophores follow nearby melanophores. These interactions are parameterised by the data observed by Yamanaka and Kondo [42]. Woolley *et al* predicted that without any accompanying mechanism, other than chase-run, the pattern will always be unstable: that is, melanophores will always by chance, move into regions densely populated by xanthophores and as a result the xanthophores in that region will then follow the melanophores once more, destabilising the pattern. Indeed, they corroborate their claims with stochastic agent-based simulations, suggesting that additional constraints on cell behaviour are required for pattern formation.

They find that even altering the parameters to match the speed and direction estimates for melanophore-xanthophore interactions in the *jaguar* and *leopard* mutant

experiments [42] did not affect the collective movement of cells significantly, suggesting that the chase-run mechanism is not sufficient for generating the pattern alone.

In 2015, Painter *et al.* generated similar results using a different model in a more comprehensive theoretical study [24]. They produced a more general system of one-dimensional non-local integro-differential equations as a framework for modelling cell populations in which there is direct contact (one to one signalling) between cells in the form of attraction and/ or repulsion. Using this two-species non-local model, they show that the 'chase-run' mechanism, modelled as attraction of xanthophores to melanophores and repulsion of melanophores by xanthophores with no homotypic interactions, can neither generate nor maintain the striped pattern seen in zebrafish. Therefore, like Woolley *et al*, the authors concluded that chase-run mechanisms alone are not sufficient to produce nor maintain a pattern. However, unlike Wooley *et al*, they extend their investigation in order to determine what minimal additional interactions are sufficient for stripe formation with chase and run. They find that only with additional positive homotypic interactions between xanthophores could the system create the conventional wild-type zebrafish pattern with a run-chase mechanism although they acknowledge that there is no evidence for this *in vivo*.

Both studies show that whilst a chase-run mechanism may be an integral part of pattern formation, alone it can neither generate nor maintain a striped pattern and thus must act as a single component in a more complex system.

4.4.4 Summary of component-level models

It is clear from the work summarised in this section that whilst component-level modelling cannot answer questions about an overarching system, that they can be important for determining the sufficiency of biologically determined interactions to generate patterns. Indeed in all of the zebrafish related models we will consider, many of the results that have lead to hypotheses relating more generally to pattern formation have built upon component-level models [13, 24, 17] (see Section 4.6).

Component-level models are typically reactive to new findings and in general aim to test hypotheses that are difficult to test experimentally. For example, following findings that melanophore and xanthophores cells exhibit a chase-run phenomena (Yamanaka *et al* in 2014 [42]), two independent groups performed mathematical analysis into whether a chase-run mechanism could be sufficient for stripe formation [13, 24]. Each group subsequently determined that it was not, however, oftentimes it inspired further investigation. For example Painter *et al* [24] determined that the minimal additional requirement for stripe formation with chase-run (considering only movement) is attraction of melanophores.

A limitation of these models is that unlike comprehensive models, besides being able to determine whether a particular feature is important, due to their simplicity, they rarely provide new avenues of research specific enough to be testable experimentally.

4.5. Comprehensive models

Comprehensive models for zebrafish pattern formation are those that incorporate as much biological information pertaining to pattern formation as possible with reasonable justifications for any variables or mechanisms that are ignored. These models can be divided into discrete and continuous models.

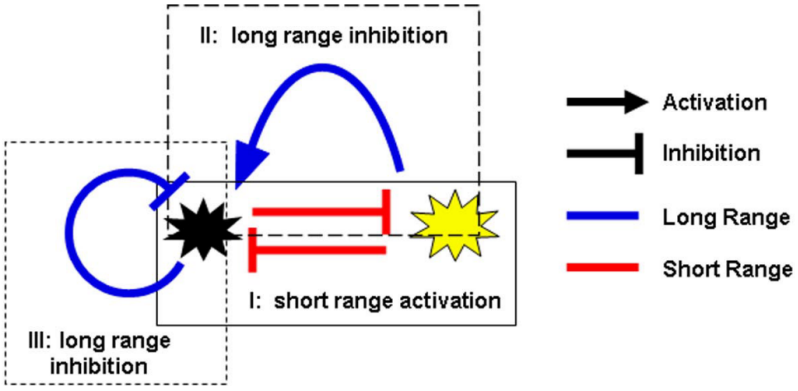


Figure 4-1: ***Mechanism underlying all proposed Turing reaction-diffusion models for zebrafish stripe formation later than 2009 (recreated from Nakamasu et al [18]).*** *Blue lines represent the long-range interactions deduced from ablation experiments. Red lines represent the short-range interactions deduced from ablation experiments. I) Melanophores and xanthophore inhibit the birth and survival each other in the short range. II) Melanophores inhibit the survival other melanophores in the long range. III) Xanthophores promote the birth and survival melanophores in the long range.*

4.5.1 Continuous models

Continuous mathematical models of zebrafish pigment pattern formation represent cells as continuous densities rather than discrete individuals. Thus far, comprehensive continuous models only consider the Turing mechanism of pattern formation.

Zebrafish pattern formation as a Turing mechanism

The original Turing model [91] is a reaction-diffusion model with special properties. A reaction-diffusion model is a simple system typically comprising two diffusing morphogens that interact with each other whilst diffusing at different rates and a Turing model is a reaction-diffusion model in which diffusion drives the generation of spatial patterns. In 1952, Turing demonstrated that this simple model could generate a wide range of patterns observed in nature.

Inspired by the beautiful simplicity of this postulated pattern forming process, many mathematical and biological investigations have since been undertaken to determine whether this approach accurately represents patterns in nature. Indeed, using computational simulations, various theoretical studies have since shown that this type of model could generate a wide range of pigment patterns observed on animal skin [92, 93, 94, 95] although similarity does not necessarily imply accuracy nor biological relevance.

Recent studies have focussed on the stripe pattern of zebrafish as a model organism for the study of pattern formation, many of which have applied a Turing model approach [16, 20, 18]. In particular, based on the erroneous assumption that iridophores are not necessary for stripe formation (since they do not appear in the striped fins, appear in both the stripe and interstripe region and have qualitatively similar interaction with melanophores as xanthophores), many Turing-type reaction-diffusion models have been proposed describing stripe formation based on interactions between melanophores and xanthophores only. In the next few subsections we will summarise all Turing models of zebrafish stripe formation.

Asai *et al* 1999

The first Turing reaction-diffusion model for zebrafish pattern formation was presented in 1999 by Asai *et al* [21]. In this paper the authors develop a simple reaction-diffusion model representing two unknown activator and inhibitor chemicals that they suggest could provide a chemical prepattern for the differentiation of melanophore and xanthophore cells. Using the model they show that they can reproduce both WT striped patterns as well as the spotted *leo* pattern. The patterns generated by this model are very convincing, however, since 1999, the idea of a chemical pre-pattern in the case of zebrafish has since been ruled out due to observations that mutants lacking a cell type do not exhibit stripe patterns [46].

Yamaguchi *et al* 2007

In 2007, Yamaguchi *et al* [22] develop the first turing model of zebrafish pattern formation based on cell-level observations. In particular, their Turing model was based on observations of melanophore and xanthophore regeneration post *in vivo* ablation performed on an entire section of a zebrafish stripe and interstripe. They showed using a simple Turing reaction-diffusion mechanism of short range activation and long range inhibition between melanophores and xanthophores that they could recapitulate *in silico* the regeneration of stripes seen *in vivo*.

Nakamasu *et al* 2009

This was followed later by a more thorough study in 2009 by Nakamasu *et al* [18] in which they ablated smaller sections of pigment cells from regions of the stripe and interstripe during and following metamorphosis and observed the subsequent regeneration of different cell types. They observed that during regeneration the presence of xanthophores appeared to promote the survival and production of melanophores in the long-range and inhibit the production of melanophores in the short-range (Fig 4-1). This lead them to propose an informed Turing reaction-diffusion model in which melanophores and xanthophores interact using a combination of short and long-range signals in a three species (melanophore, xanthophore, long range chemical) Short range Activation, Long-range Inhibition (SALI) Turing model. The model is described by three partial differential equations. One equation describes the spatial arrangement of melanophores, one describes xanthophores (and simultaneously the short range chemical signal produced by melanophores/xanthophores), the third describes the long-range signalling molecule produced by xanthophores.

Each partial differential equation is made up of a diffusive part, describing the basic movement of the cells/chemical and a reactive part which describes the interactions between the two cell types and signalling molecules. By changing the parameter values and initial conditions, the authors show that this model is able to replicate patterns in wild-type, *obelix, leopard* and *choker* mutants as well as wild-type ablation patterns. For the wild-type condition, the system is initialised with a central horizontal stripe of xanthophores and is evolved according to the three reaction-diffusion equations. This suggests that iridophores are important for pattern orientation, as observed experimentally [39, 46, 47]. The images of the patterns that result from these simulations, when viewed alongside their experimental counterparts are qualitatively convincing, however, the model has many limitations.

Firstly, there is no evidence that information is spread between cells by a signalling molecule. In fact, subsequent studies have shown that information is instead spread by cell-cell contact or the extension of filopodia and airinemes between cells [63, 62]. Secondly, the reaction-diffusion mechanism requires that xanthophores and melanophores

must be moving. However, zebrafish pigment cells do not move quickly relative to the domain size, melanophores move at 110μm/week and xanthophores move three times more slowly [43].

Bullara and Decker 2015

Motivated by the aforementioned studies by Kondo and observations that xanthophore and melanophore cells do not exhibit large-scale movement [43], in 2015, Bullara and De Decker generated the first Turing model of melanophores and xanthophores without diffusion [20]. That is, with completely immobile melanophores and xanthophores at all time points. This chemical-reaction model is a SALI model of melanophores and xanthophores based on the Kondo ablation experiments, but without any cell movement.

The model can be simulated stochastically on a lattice or analysed deterministically by making some approximations to generate a pair of partial differential equations (PDEs). The authors show, by linear stability analysis alongside stochastic simulation, that by varying parameters this model can produce stripes and spots in various orientations. For generating horizontal stripes as seen in zebrafish, they initialise a horizontal central stripe of iridophores (implemented as an initial stripe region in which melanophores cannot differentiate using parameters within a certain range.

For generating stripes seen in zebrafish, they impose the following rules:

- random birth of xanthophores,
- long-range promotion of melanophore birth by xanthophores,
- short-range inhibition of xanthophore birth by melanophores and melanophore birth by xanthophores with parameters in a certain range.

The authors find that the wavelength of the pattern is twice the length of the long-range interaction suggesting that, if this mechanism is correct, then the width of the stripe is influenced by length of the long-range interactions between melanophores and xanthophores. The limitations of this approach are that, whilst the patterns it produces are convincing at large time scales (approximately at 1,000,000 steps which could be many weeks in real time), the evolutions of the stripe pattern (seen at earlier time points) are not consistent with the literature (compare Fig. 4-2 and Fig. 3-2(C)). For example, in the stochastic simulations, melanophores and xanthophores repeatedly aggregate and rearrange until finally reaching a stable pattern of stripes. This is inconsistent with the observed biological mechanism, which shows clear consecutive formation of stripes, starting with the original central stripe followed by the nearby dorsal and ventral stripes [31], suggesting that whilst SALI could play an important role in pattern formation other factors contribute to the timing and development of the system.

Kondo 2017

Upon finding that cell-cell interaction in the zebrafish was via cell-cell contact and not by signalling molecules as previously thought, in 2017, Kondo proposed an updated kernel-based Turing model [16] in a more theoretical study of reaction-diffusion mechanisms.

Instead of modelling signalling information being endocytosed from the cell body and transported between cells via diffusion through the surrounding medium (as might be seen by diffusion of a signalling molecule), signalling information is modelling instead as being transferred by non-local interactions. For example, by 'diffusion' of cellular

(a) (b) (c) (d) (e)

Figure 4-2: *Evolution of pattern formation using the model by Bullara et al [20]. The images show the spatial domain at a) 60,000, b) 150,000, c) 7,000,000 d) 46,000,000 e) 100,000,000 Monte Carlo steps. Yellow and black space represents space occupied by xanthophores and melanophores respectively.*

projections such as filopodia or airinemes. The model is made up of two partial differential equations describing the spatial arrangement of two cell species over time. Cell move down a diffusion gradient and interact via short range and long range cellular projections which send signals for death or differentiation of different cell types. This is modelled mathematically by the addition of two Gaussian functions centred at two points some distance away from the cell body, representing the length of the short and long range cell projections. For SALI-type interactions, of the two Gaussian functions, one is typically an activator function $A(x)$ and the other an inhibitor function $I(x)$. The activator function is positive and promotes the differentiation of the other cell type, whilst the inhibitor function is negative and therefore promotes the death of the other cell type.

Upon varying the strength and location of the centre of the activator and inhibitor functions, Kondo demonstrates the ability to produce many different patterns including stripes, spots and even nested patterns comprised of two patterning modes superimposed on each other, showing that reaction-diffusion patterns can still arise without assuming information transportation via signalling molecules.

Whilst these models provide some insight to the mechanism involved in generating stripes, the hypothesis that zebrafish pigment pattern formation is a consequence of a Turing mechanism has been challenged recently by the findings that iridophores are necessary for formation of stripes on the body of the fish [46]. A further criticism of Turing mechanisms is that these models rely heavily on the results of the ablation experiments by Kondo in 2009 [18] and neglect other observations.

A summary of Turing models

Turing reaction-diffusion mechanisms are frequently employed by mathematicians to explain a variety of pattern formation mechanisms due to their elegant mathematical simplicity and their ability to qualitatively replicate patterns. However, it seems likely that these models are too simplistic to explain zebrafish pigmentation patterns. Furthermore, since the SALI assumption are made from observations in ablation experiments post metamorphosis in fully developed stripes, it leads to questions whether these interactions are involved during pattern formation or simply in pattern regeneration and maintenance. The benefits and limitations of each of the models are summarised in Table 4.1.

Paper	Mechanism	Model type	Pros	Limitations
Asai *et al* (1999) [21]	Long-range activation short-range inhibition mediated by long-range/short-range signalling molecules.	Continuous PDE.	Able to replicate patterns for wild type and *leo* mutant.	Model not based on evidence, only pattern observation.
Yamaguchi *et al* (2007) [22]	Long-range activation short-range inhibition mediated by long-range/short-range signalling molecules.	Continuous PDE.	Able to replicate adult patterns for wild type as well as able to replicate regeneration experiments.	Model not based on evidence, only pattern observation.
Nakamasu *et al.* (2009) [18]	Long-range activation short-range inhibition facilitated by long-range/short-range signalling molecules.	Continuous PDE.	Able to replicate pattern formation for wild type and some mutant zebrafish.	Incorrect assumption of cell-cell communication though signalling molecules.
Bullara *et al.* (2015) [20]	Long-range activation short-range inhibition by melanophores and xanthophores.	Stochastic ABM and corresponding continuous PDE.	Able to produce striping pattern.	Dynamics of pattern formation is not observed in zebrafish pattern formation.
Kondo (2017) [16]	Turing model in which communication is facilitated by interactions by cell projections.	Continuous PDE.	Able to produce a wide variety of biological patterns.	Implicitly relies on diffusion of cells as mode of transport (cells do not move substantially in reality).

Table 4.1: *Summary of Turing reaction-diffusion models for zebrafish pattern formation in order of publication.*

4.5.2 Discrete models

Discrete mathematical models of zebrafish pigment pattern formation represent cells as discrete agents instead of as continuum densities. These models can be further split into on-lattice model which model cells distributed on a lattice and off-lattice models which allow for continuous movement of cells in space. We summarise the findings from these models in the following subsections. As the findings are more complex than in the continuum case we label each subsection with the main author.

Moreira *et al* 2005

The first on-lattice agent-based mathematical model of zebrafish pigment pattern development was presented in 2005 by Moreira *et al* [25].

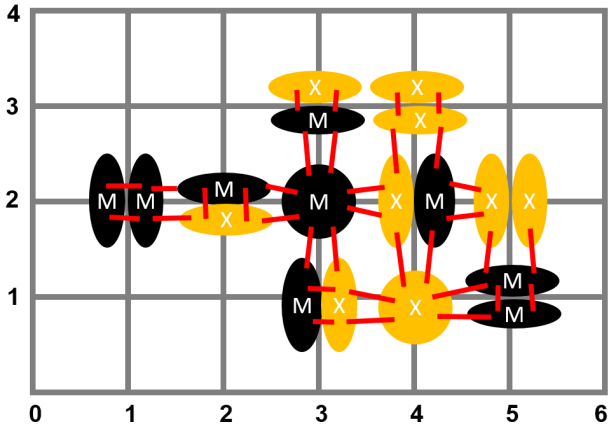


Figure 4-3: ***The neighbourhood of a cell in the model of Moreira et al.*** *Recreated from Moreira et al [25]. Diagram showing the bindings between chromatophores considering a von Neumann neighbourhood, which reflects different contact surfaces between adjacent cells, depending on the spatial configuratons of the nodes. Black and yellow cells marked with M and X represent melanophores and xanthophores respectively. The eleven possible node configurations i.e. the positioning of the chromatophores in the node are depicted in nodes: (1,1) no chromatophores; (3,3) single melanophore; (4,1) single xanthophore; (5,2) and (4,3) two xanthophores; (1,2),(5,1) two melanophores; (2,2),(3,1), (3,3) and (4,2) a melanophore and a xanthophore.*

Inspired by genetic studies suggesting that adult pigment pattern formation and maintenance in zebrafish resulted from local interactions between cells and not from long-range signalling (as had been conjectured previously [21]) Moreira *et al.* developed a two-species melanophore-xanthophore cellular-automaton model regulated entirely by local cell-cell interactions.

They modelled a monolayer of xanthophore and melanophore cells, as a two-dimensional domain, divided into a square lattice of which, each node can contain up to two chromatophores, (melanophores and xanthophores only), in one of four possible spatial arrangements (Fig. 4-3). During the simulation, at any given time, each cell has a corresponding age (days since 'birth') and a corresponding energy relating to adhesive properties between neighbouring cells wherein a lower energy state is more desireable than a higher energy state. For example, adhesive properties may be chosen such that xanthophores and melanophores are strongly adhesive, in which case the energy between cells of this type would be low. Adhesive properties may be chosen such that xanthophores are highly repelled by other xanthophores and hence the energy between these cell types would be high. The total energy of any given cell in a particular position at a given time is the total energy between the cell and all of its' neighbouring cells. These energies are used to incorporate local interactions between different cell types and to determine the movement of cells. In this model a cell will always prefer to be the lowest energy state possible.

The model is initialised to represent the early larval pattern formed in the embryo of a fish aged approximately 6 dpf: as four rows of melanophore stripes equidistance apart. In particular, there are two rows of melanophores representing the dorsal and ventral melanophore stripes and two adjacent rows of melanophores representing the central stripes. Unfortunately, this interpretation of the early pattern, which was not uncommon for the very early models, was a misinterpretation of the pattern. In particular, the melanophore stripes they describe are actually much deeper in the hypodermis and would not be at skin level. In particular, the two central stripes are actually associated

with the neural tube [45].

The domain is then updated ten times each simulated day. At each time step, the new configuration of the system results from updating chromatophore age, birth, death and movement. Chromatophore birth is determined by choosing a random subset of nodes and assigning a temporary 'stem cell' to them. Birth of new pigment cells by proliferation of stem cells and differentiation is then described using two competing hypotheses:

1. stem cells originate near stripes which differentiate into xanthophores initially and then later, whatever is locally in abundance: a 'community effect';

2. stem cells are uniformly distributed throughout the domain and give rise first to xanthophores, and subsequently by a community effect with a bias towards xanthophores in unoccupied regions of the domain.

All chromatophores are considered in turn for death and subsequently upon survival, movement at each time step. For each cell, a death event is successful if

1. the cell has reached their mean lifetime and

2. the cell is isolated from cells of the same type.

This is another form of community effect. If one or neither of these criteria are true, the chromatophore survives and instead the chromatophore attempts movement. In this case, a direction is chosen uniformly at random between available (unoccupied) local sites.

Movement into this space is successful (and the cell moves from its original space into the new space) if the energy of the cell in the new configuration is less than or equal to the energy in its original position, otherwise the chromatophores remains in its original position. Growth of the domain in the horizontal direction is incorporated by introduction of new rows at a rate of 35μm/day from 6-23 dpf; 130μm/day from 23-38 dpf. Growth in the horizontal direction is neglected.

The authors show that only the second hypothesis for chromatophore birth can reproduce striping. The authors find that by varying the adhesive coefficients, they can produce qualitatively similar patterns to those seen in wild-type *D. rerio* and in *leopard* and *obelix* mutants.

Since the study was conducted in 2005, the experimental findings it draws upon are restricted to those generated before this time, placing an inherent limitation on the realism of the model. One of the main restrictions of this model is the assumption that melanophores and xanthophores are attracted to each other in the wild type case, which we now know not to be true [42]. Furthermore, there are an abundance of unknown parameters (more than five) in the model which are now known and which do not fit the models original assumptions.

Nevertheless, the model generates insight into how local (microscale) interactions can generate patterns in the macroscale, in particular, how changes to local interactions can have detrimental impacts on the type of pattern, for example from stripes to spots as seen in the *leopard* case. It was also the first mathematical model of zebrafish to identify community effect as a feasible patterning mechanism.

Volkening *et al* 2015

In 2015, Volkening *et al* presented a comprehensive two-species mathematical model of zebrafish pigment pattern development and regeneration [19]. The body of the

fish, in the monolayer of xanthophore and melanophore cells where striping occurs, is represented by a two-dimensional domain, containing melanophores and xanthophores only represented as point particles. Iridophores are also incorporated implicitly.

Cells interact through: migration, governed by short-range cell-cell interactions; differentiation and death, governed by long-range cell-cell interactions on a growing (and as a special case as a static) two-dimensional rectangular domain. The model is initialised at approximately 7.2mm SSL, as four larval stripes of melanophores, and a central stripe of xanthophores. As in the model of Moreira *et al* [25] this is a misinterpretation of the larval pattern. Movement of each melanophore and xanthophore is governed by ordinary differential equations (ODEs). The functional form of these ODEs is dependent on the cells' contact-range (36μm) and short-range interactions (82μm) with other cells in the domain. Cell-cell interactions are modelled using Morse potentials. The potentials are determined by the following cell-cell interactions observed from both *in vitro* and *in vivo* experiments:

1. attraction of xanthophores by melanophores at short-range;
2. contact repulsion and short-range attraction of xanthophores by melanophores;
3. contact attraction and weak short-range repulsion of xanthophores by melanophores (variations of a chase and run mechanism observed between melanophores and xanthophores studied *in vitro* by Yamanaka and Kondo 2014 [42] are investigated); and
4. repulsion of xanthophores by melanophores at short-range as observed on the body of the fish by Frohnhofer *et al* 2013 [46].

Volkening *et al* found that the chase-run mechanisms (3) did not produce patterns comparable to wild-type but that mutual repulsion between cells of types could. Furthermore, they found that this mutual repulsion between melanophores and xanthophores could be responsible for maintaining the stripe and interstripe boundary.

Mutual repulsion is considered as the chosen patterning mechanism going forward. Death and differentiation rules are based on the findings by Nakamasu *et al* [18] that xanthophores in adjacent stripes promote the differentiation and survival of melanophores at least half a stripe width away.

At the end of each simulated day, a random subset of cells is chosen for assessment of the death condition. If a chosen cell is within short-range distance (40 μ m) of like-neighbours and at a long-range distance (a stripe width distance) from the opposing cell type then the focal cell survives, otherwise it dies. Similarly, a random set of unoccupied positions are chosen for possible cell differentation. A cell differentiates in this space if there are is a high proportion of same-type to opposite-type cells in the short-range and a high proportion of opposite-type to same-type cells in the long-range. There is also a small probability a xanthophore cell can differentiate in this space regardless of the local environment. Growth is modelled by uniform stretching wherein cells are pulled in both horizontal and vertical directions by 130μm each day.

Finally, as a proxy for the initial iridophore stripe and sequential spreading of iridophores, where iridophore aggregation for each of the stripes occurs, a line of xanthophores appear in corresponding regions at different stages, based on findings that xanthophores are highly associated with iridophores. By changing the time when subsequent xanthophore stripes appear (as a proxy for iridophores) the authors investigate the role of iridophores implicitly.

The authors report that 'iridophores' guide pattern directionality, as stripes form more robustly when sequential spreading of iridophores is implicitly incorporated around

9.6mm SSL or 4 days into the simulation. In contrast, bridges form between stripes when the sequential spreading of iridophores appear any earlier. They are able to replicate ablation experiments by initialising the domain with a region unpopulated by cells. Whilst their direction is perturbed in these experiments, stripes have the same width as in real fish. The authors show that they are also able replicate pattern formation in *choker* mutants which lack a horizontal myoseptum. By comparing growing and static domains they find that growth in both directions may widen (in the vertical axis) stripes and interstripes, allowing for shorter long-range interactions. By allowing model simulation with just growth or just movement, they find that growth plays a role in maintaining stripe width (in the vertical axis), whilst cell movement, attractive, repulsive dynamics maintain stripe-interstripe boundary. The authors find that the length of the long range interactions determining differentiation of new cells specifies stripe pattern width, and that migration of cells (exhibiting mutual repulsion) appears to maintain it.

Volkening *et al* 2018

In 2018, Volkening *et al* built upon their original model [19], by incorporating neglected cell-type iridophore explicitly in their model [26]. The body of the fish is represented as three two-dimensional domains corresponding to the three layers of pigments cells in the real fish: melanophore, xanthophore and S-iridophore layers (Fig. 3-2(A)). Pigment cells are modelled as individual agents of five different types: melanophores, loose and dense xanthophores and loose and dense iridophores, which occupy space on their allocated layer only.

The model is initialised at 7.2 mm SSL as melanophores and loose xanthophores scattered uniformly at random at a low density and a central stripe of iridophores are positioned on the domain to represent the iridophores that appear along the horizontal myoseptum. This is an improvement on their previous interpretation of the initial pattern and represents more accurately, the cells on the skin at this time.

In a manner consistent with their previous model, cells self-organise through migration (determined by ODEs incorporating short range repulsion and attraction rules), as well as differentiation and death (governed by short and long range interactions). Growth is implemented as in Volkening *et al* 2015 [19].

Unlike the previous model, which did not distinguish loose and dense xanthophores, nor incorporate cell type iridophore, loose and dense shape transitions are also incorporated, governed by cell-cell interactions. Loose and dense shape transitions for xanthophores are informed by the literature. In contrast, loose and dense iridophore shape changes are determined by investigating the three missing cell mutants: *pfe*, *nac* and *shd*. Determining the interactions involved in determining iridophore shape changes is the major finding of this paper. In particular, based on their observations, the authors deduce that iridophore shape changes must be guided by cues from melanophores and dense xanthophores. Moreover, they deduce minimal rules for shape changes elicited from melanophores and dense xanthophores, that can reproduce wild-type and missing cell mutant patterns.

The model is updated every 24 hours of simulated time. In particular, the domain size grows as in Volkening *et al* 2015. Cell positions are updated according to the ODEs for migration. Finally, cells on the domain are evaluated for possible death (melanophores only) or shape transitions (xanthophores and iridophores only).

The authors demonstrate that, using the full set of rules, including their derived rules for iridophore shape changes, not only can they reproduce wild-type striped pat-

terns, but that they are also able to reproduce missing cell mutant patterns *shd, nac* and *pfe* as well as double mutants *pfe;nac*, *pfe;shd* and *nac;shd*. They further demonstrate the capability of the model by also replicating mutants *choker* and *puma* as well as the results of some ablation experiments. By increasing the strength of signalling required for successful iridophore shape changes, the authors show that the wild-type and missing cell mutant patterns are robust to noise.

Finally, having validated their model, the authors use the model to address *Danio* evolution. They show, that by removing some of the requirements for shape transitions of iridophores that they can reproduce some of the patterns observed in close relatives *Danio albolineatus* and *Danio margaritatus*. They hypothesise that these relatives differ from zebrafish due to the loss of iridophore shape signalling.

There is no doubt that these studies by Volkening *et al* [19, 26] make some excellent contributions to the understanding of zebrafish stripe formation. By incorporating all of the known cell-cell interactions, the authors are able to analyse key aspects of zebrafish stripe formation and assess the validity of various hypotheses.

4.5.3 Summary of comprehensive models

A major benefit of comprehensive modelling is that they can be used to investigate questions specific to the system for example; gene function [26]. A limitation of comprehensive models is that specificity of the model to the zebrafish stripe formation often makes these findings not generalisable to other patterns.

In the context of zebrafish stripe formation, comprehensive mathematical models have improved dramatically and are more useful than ever. Initially simplistic two-species models with little complexity such as Moreira *et al*'s individual based model [25] as well the original Turing models of melanophores and xanthophores [22], have evolved to detailed five species models with over a 100 different cell-cell interactions [26]. This is to some extent a reflection of the advances made in computational power and the ability of such models to be simulated and therefore postulated in the first place, but mostly it is a reflection of the biological evidence now available. Experimental studies over the last 15 years have gradually revealed the underlying mechanisms of pigment pattern formation in zebrafish as a very detailed and complex system.

The future of comprehensive models of zebrafish will be in aiding further investigation into the remaining questions surrounding pattern formation, such as gene function and *Danio* evolution as seen in Chapter 3.

4.6. Conclusion

Zebrafish have become a model organism for pattern formation and mathematical modelling is a powerful assistive tool for understanding zebrafish stripe formation. In this review we have described the relevant mathematical models of zebrafish pattern formation along with their implications and limitations. A summary of these models is given in Table 4.2. It is clear that, not only have these mathematical models contributed to the understanding of zebrafish stripe formation, but many have also contributed to the understanding of pattern formation more generally.

In summary, the results of these models have lead to the following hypotheses relating to zebrafish stripe formation:

1. The chase-run mechanism observed between melanophores and xanthophores [42] neither produces nor maintains stripes [13, 24].

2. Spreading and aggregation of iridophores may guide pattern directionality [19].
3. Growth may maintain stripe width [19].
4. Cell movement with attractive and repulsive dynamics maintains stripe and interstripe boundaries [19].
5. Differentiation and death may specify stripe pattern width [19].
6. The length of melanophore and xanthophore interactions may determine the width of stripes [20].
7. Iridophores may orient the stripes [18, 19].
8. Iridophores are crucial for pattern development [26].
9. Potential rules for iridophore shape changes [26].
10. Zebrafish may be related to *D. alboneatus* and *D. margariatus* through a loss of iridophore shape signalling [26].

Additionally, the results of these models have lead to the following hypotheses about more general pattern formation:

1. Short range interaction alone between two species can produce patterns [25].
2. Chase-run between two species is not sufficient for pattern formation. [13, 24, 19].
3. Community-effects can generate stripes [17, 25].
4. Strong heterotypic repulsion with weak homotypic repulsion between two cell species can generate stripes [23].
5. A reaction-diffusion model for which information is passed by cell-cell signalling instead of via a signalling molecule can still generate patterns [16].

The primary role of mathematical modelling of biological processes is its ability to test and verify experimental hypotheses, as well as to make predictions which can indicate appropriate experimental directions. In this review we summarised both component-level and comprehensive models of zebrafish pattern formation. Whilst both techniques have their benefits and limitations, it is clear from this review that comprehensive models have more scope for testing hypotheses and generating testable hypotheses relating to the specific development of zebrafish stripe formation. Meanwhile component-level models are useful for determining whether individual elements of a system are sufficient for pattern development in a more general sense but are less useful for system-specific predictions. For this reason, component-level models are generally not ideally suited to generating testable predictions regarding zebrafish pattern formation.

Of the comprehensive models, there is further dichotomy between the types of models used to understand zebrafish stripe formation. On one side of the divide are continuum Turing models. Turing models are multi-species reaction-diffusion models that postulate diffusion-driven instability leading to pattern formation. In zebrafish Turing models, xanthophores and melanophores are typically the two diffusing species and the reactions are short and long range birth and death signals between melanophores and xanthophores deduced from results of ablation experiments. Aside from this, no other biological findings are incorporated. A great benefit of these models are their

ability to generate not only stripes but a whole range of zebrafish mutant patterns [71] – theoretically almost any pattern can be generated by tuning components of the core Turing mechanism. However, a significant limitation is that, despite recent improvements [16], parameters are often difficult to link back to the real system. Particularly damaging to the use of Turing theory is the fact that cells do not migrate quickly [43] – contradicting a key assumption of diffusion-driven instability theory – as well as the fact that iridophores have been shown to be pertinent for stripe formation [12] and are not incorporated explicitly in the two species models.

The other type of comprehensive models are discrete models. These models represent cells as individual agents that move and interact in a way that is directly relatable to the real system. Over time, and due to the abundance of zebrafish experimental studies, these models have become more complex, incorporating over a hundred different cell-cell interactions. A significant benefit is that, due to their direct relatability to the real system, these models can be used to make predictions and test hypotheses that are specific to zebrafish development, such as investigating the function of different genes or the impact of growth.

In conclusion, there have been an abundance of mathematical investigations into zebrafish stripe formation. So far most of these studies have focussed on the sufficiency of the biological rules in determining patterns and few have been used for predictive investigations. Since many of the biological mechanisms for zebrafish pattern formation have now been outlined, focus will be turned to the molecular details that remain unknown. In this context, developing mathematical models capable of investigating such intricate phenomena represents an important avenue for future research.

Paper	Summary	Main findings
Moreira *et al* (2005) [25]	On-lattice model of pattern formation. Melanophores and xanthophore move according to attraction between other cell species. Community effect model for differentiation and death.	• Short-range interactions alone can produce patterns. • Community-effect differentiation and death can produce striped patterns.
Caicedo-Carvajal *et al* (2008) [23]	Component-level stochastic two species model of heterotypic and homotypic attraction and repulsion.	• Strong heterotypic repulsion with weak homotypic repulsion between two cell species can generate patterns similar to those seen in zebrafish stripes (*i.e.* alternating stripes of cells with space between them).
Nakamasu *et al* 2009 [18]	Turing reaction-diffusion mechanism of short-range activation and long-range inhibition.	• Short-range activation and long-range inhibition can produce striping similar to that seen in zebrafish. • Iridophores may align stripes horizontally.
Bloomfield *et al* (2011) [17]	Integro-differential equation for community-effect between two species.	• Community effect can generate stripes, provided the death rate of the two species is equal and the transition from random differentiation to proximity differentiation is steep.
Wooley *et al* (2014) [13]	Run and tumble model of melanophores and xanthophores with chase-run mechanism.	• Chase-run mechanism between melanophores and xanthophore neither produces nor maintains stripes.
Painter *et al* (2015) [24]	Integro-differential equation for chase-run of melanophores and xanthophores in 1D.	• Chase-run mechanism between melanophores and xanthophores neither produces nor maintains stripes. • The minimal additional requirement for stripes with chase-run is attraction of melanophores.

Volkening *et al* (2015) [19]	Off-lattice model for melanophore-xanthophore interactions.	• Chase-run is not sufficient for stripe formation. • Mutual repulsion, however, is, and could maintain stripe and interstripe boundaries • Spreading/aggregation of iridophores guide pattern directionality. • Growth may widen stripes and interstripes, allowing for shorter long-range interactions. • Growth appears to maintain stripe width. • Differentiation and death (short-range activation and long-range inhibition) may specify stripe pattern width, • Migration of cells may maintain stripe and interstripe boundary.
Bullara *et al* (2015) [20]	Stochastic simulation and corresponding deterministic PDEs for long-range inhibition and short-range activation without movement.	• Movement is not required for the Turing short-range activation and long-range inhibition stripe formation. • The length of long-range interactions between melanophores and xanthophores may determine the width of stripes.
Kondo (2017) [16]	An updated kernel-based Turing model.	• Short-range activation and long-range inhibition can produce striping similar to that seen in zebrafish.
Volkening *et al* (2018) [26]	Off-lattice model for melanophore, xanthophore and iridophore interactions leading to pattern formation.	• Iridophores are crucial for pigment pattern formation. • Derive rules for iridophore shape changes. • The model suggests that *Danio alboneatus* and *Danio margaritatus* could be related to zebrafish through a loss of iridophore-shape signalling.

Table 4.2: *Table summarising the mathematical models in order of publication and their contributions to understanding zebrafish pattern formation.*

4.7. Chapter Conclusion

This chapter provides an overview of the all current mathematical models of zebrafish stripe formation along with their major insights into the pattern forming process. The chapter gives an overview of the two types of modelling: component-based and comprehensive modelling in the context of stripe formation as well as their benefits and pitfalls.

In the context of the thesis, the chapter provides motivation for the comprehensive modelling approach to zebrafish stripe formation which we present in Chapter 6. In particular we use a stochastic on-lattice volume exclusion model, wherein cells are represented as agents occupying lattice spaces that interact in a manner consistent with the literature.

In the next chapter we turn our attention towards a means of validating an on-lattice mathematical model of stripe formation. To this end, we generate a statistical tool that can both identify and quantify spatial patterns in discrete domains. This will enable us to compare our model simulations with the corresponding fish patterns.

Chapter 5

Pair correlation functions for identifying spatial correlation in discrete domains

5.1. Overview

Motivated by the need to be able to quantify spatial patterns on the lattice domain of the comprehensive mathematical model that we build in Chapter 6, this chapter is dedicated to the problem of identifying and subsequently quantifying spatial correlation in discrete domains. This chapter comprises the paper written by myself (Jennifer Owen), Enrico Gavagnin and Christian Yates in 2018 as published in Physical Review E [29].

Spatial correlation is a dominant feature of many biological and physical systems [96, 97, 98, 99, 100, 101, 102, 103, 104]. The adult zebrafish pigment pattern is a classic example of spatially correlated pigment cells. Xanthophores, for example, are spatially correlated at short range (neighbouring) distances and anti-correlated in long range (stripe width distance apart).

In this paper, we design a series of novel pairwise correlation functions (PCFs), for multi-agent systems in discrete domains, with the aim of using this function to quantify patterns generated using our discrete mathematical model of zebrafish stripe formation in Chapter 6. Prior to this research there had been exactly two on-lattice PCFs defined. The first is the Annular PCF, which applies the classic off-lattice PCF to lattice-based systems. In the Annular PCF, given some small positive δ, the number of agents at distance m from a focal agent is defined as the number of agents whose centres lie in the annulus $(m-\delta,m]$, where distance is defined by the Euclidean metric. A limitation of this method is that, whilst the normalisation is a good approximation for a continuous domain, it is poor in the case of a discrete domain, thus the PCF is not correctly calibrated. The second is the on-lattice PCF, presented by Binder and Simpson [105] which we refer to as the *Rectilinear* PCF. Whilst their approach correctly identifies the spatial correlation in many examples, due to an anisotropic definition of distance, spatial structures which are biased in either Cartesian directions can remain unidentified by this PCF. Our functions improve upon previous PCFs and are extended to different metrics and higher dimensions.

5.2. Outline of the paper

Section I of the paper provides motivation of the importance of studying spatial correlation as well as a short literature review. In Section II we evaluate preexisting PCFs for discrete domains and discuss their limitations, providing motivation for our study. In Section III we introduce our first two new PCFs for square lattices - the square taxi-cab PCF and square uniform PCF. We test the new PCFs using examples of patterns taken from nature and present our findings in Section IV. Extensions to more general discrete domains including triangular, hexagonal and cubic are discussed in Sections V and VI. We conclude with a short discussion in Section VII. The supplementary material contains derivations that were omitted in the main text, as well as a summary of the normalisations of all the PCFs defined in the paper.

Appendix B: Statement of Authorship

This declaration concerns the article entitled:			
Pair correlation functions for identifying spatial correlation in discrete domains			
Publication status (tick one) **Draft manuscript** [ ] **Submitted** [ ] **In review** [ ] **Accepted** [ ] **Published** [X]			
Publication details (reference)	Journal: Physical Review E 97, 062104 Authors: Enrico Gavagnin, Jennifer P. Owen and Christian A. Yates		
Copyright status (tick the appropriate statement) I hold the copyright for this material [ ] Copyright is retained by the publisher, but I have been given permission to replicate the material here [X]			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	The first two authors contributed equally to the idea formulation,all the calculations and simulations. (50%). All authors contributed equally to the presentationof the content (33%).		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed	Jennifer Owen	**Date**	07/06/2021

Pair correlation functions for identifying spatial correlation in discrete domains

Enrico Gavagnin,[*,†] Jennifer P. Owen,[†,‡] and Christian A. Yates
Centre for Mathematical Biology, University of Bath, Claverton Down, Bath, BA2 7AY, United Kingdom

(Received 12 January 2018; published 4 June 2018)

Identifying and quantifying spatial correlation are important aspects of studying the collective behavior of multiagent systems. Pair correlation functions (PCFs) are powerful statistical tools that can provide qualitative and quantitative information about correlation between pairs of agents. Despite the numerous PCFs defined for off-lattice domains, only a few recent studies have considered a PCF for discrete domains. Our work extends the study of spatial correlation in discrete domains by defining a new set of PCFs using two natural and intuitive definitions of distance for a square lattice: the taxicab and uniform metric. We show how these PCFs improve upon previous attempts and compare between the quantitative data acquired. We also extend our definitions of the PCF to other types of regular tessellation that have not been studied before, including hexagonal, triangular, and cuboidal. Finally, we provide a comprehensive PCF for any tessellation and metric, allowing investigation of spatial correlation in irregular lattices for which recognizing correlation is less intuitive.

DOI: 10.1103/PhysRevE.97.062104

I. INTRODUCTION

A system of agents is considered in a state of spatial correlation if, given any agent in the system, the likelihood that there are other agents at a certain, close distance, is either increased or decreased with respect to the situation in which the agents are distributed uniformly at random. Spatial correlation is a dominant feature of many biological and physical systems [1–12]. For example, in cell biology, spatial correlation can be seen in the form of patterns on animal fur or fish skin [1,3,13]. In a clinical setting, cell aggregation is a characteristic feature of melanoma and its identification is essential for early diagnosis and effective therapy [14,15]. Resource competition in ecology can lead to spatial correlation in the form of segregation, for example, in ant nest displacement in a competitive environment [16]. In epidemiology, spatial correlation can be observed in the occurrence of disease across different geographical regions [17].

The same spatial configuration can have different origins. For example, spatial aggregation in cell biology can be caused by a result of cell-to-cell adhesion [8], external signals, as in *chemotaxis* [5,6], or even slime following [7]. Alternatively, cells may form clusters during development due to a combination of a high proliferation rate and a low movement rate [9]. Given a system exhibiting spatial correlation, one may hypothesize an underlying mechanism responsible for these properties. These assumptions may form the basis of a mathematical model that can be simulated for the purpose of testing. Quantifying spatial correlation in both the simulation and observed experimental data can be a way to connect these studies and to validate or disprove such a theory. As a result, a great number of statistical tools have been developed in the past decade to analyze and measure spatial correlation [12,18–23]. Among the most popular are pair correlation functions (PCFs) [11,12,20–22,24–28] and the fast Fourier transform (FFT) [29,30]. In this paper we focus our attention on the study of spatial correlation using PCFs.

Given a system of agents, a PCF determines whether pairs of agents are more or less likely to be found with a given separation than in the situation in which the agents are positioned uniformly at random in the domain. A PCF is considered effective if it fits two main criteria. First, the PCF distance metric should be well-defined, but most importantly be readily interpretable in the context of the system considered. This criteria is essential so that in the case of correlation (aggregation or segregation), the PCF can be used to obtain more details about the spatial configuration. For example, if the system exhibits aggregation, the PCF should be able to provide a measure for the average size of the clusters and their pairwise separation. Second, the PCF should be correctly calibrated. The PCF should be able to distinguish between three basic types of configurations: spatial randomness, aggregation, and segregation. For this, the PCF should be normalized correctly; i.e., the PCF should return the value unity at all pairwise distances (no correlation) when applied to a uniformly distributed set of agents. If the PCF is not normalized correctly, a spatially random set of agents may be incorrectly identified as a correlated system. This inconsistency makes PCF profiles hard to interpret.

Depending on the type of investigation, the mathematical framework can either be continuous (off-lattice) or discrete (on-lattice). The corresponding PCF has to be defined in accordance with the given framework. Despite the abundance of PCFs defined for off-lattice domains [11,26–28], only a few recent studies have defined a PCF for domains partitioned with a lattice [12,20–22]. On-lattice PCFs often assume exclusion properties, that is, that each lattice site in a domain can be occupied by at most one agent at any given time. This is consistent with typical on-lattice correlation studies, such as

[*]e.gavagnin@bath.ac.uk
[†]These authors contributed equally to this work.
[‡]j.owen@bath.ac.uk

those designed to quantify correlation in binary pixelated images, or to determine spatial correlation in exclusion processes simulated using a discrete domain.

Currently, there are two PCFs defined on-lattice. The first is a naive approach consisting of applying the classic off-lattice PCF to lattice-based systems. We refer to this from now on as the *annular* PCF. In the annular PCF, given some small positive δ, the number of agents at a distance m from a focal agent is defined as the number of agents whose centres lie in the annulus $(m-\delta, m]$, where distance is defined by the Euclidean metric. A limitation of this method is that, while the normalization is a good approximation for a continuous domain (see Sec. II for more details), it is poor in the case of a discrete domain, thus the PCF is not correctly calibrated. In more recent work, Binder and Simpson [12,21] defined a PCF specifically designed for a two-dimensional square on-lattice exclusion process, which we will refer to as the *rectilinear* PCF (see Sec. II for more details). While their approach correctly identifies the spatial correlation in many examples, due to an anisotropic definition of distance, spatial structures that are biased in either Cartesian directions can remain unidentified by this PCF. To summarize, to the best of our knowledge, a discrete isotropic PCF with correct normalization does not currently exist in the literature.

In this paper, we extend the study of pairwise spatial correlation for on-lattice exclusion processes which tackles the flaws of previous PCFs. We define isotropic PCFs for a square lattice on which distance is defined using two of the most natural and intuitive metrics for a discrete domain: the taxicab and uniform metric. We call these the *square taxicab* PCF and *square uniform* PCF, respectively, after the square lattice set up and metric type. We define them in both the nonperiodic and periodic boundary cases. Using synthetically generated data, we demonstrate that our PCF can correctly distinguish between spatial randomness, aggregation, and segregation. Furthermore, we show that it can also provide quantitative information about the structure of the system, such as approximate aggregate size or segregation distance both in the short and long scales. Moreover, we investigate how the choice of metric, uniform, or taxicab can affect this quantitative information. We demonstrate that our PCFs represent a significant improvement on previous on-lattice PCFs by showing that, first, our method is correctly calibrated (unlike the annular PCF) and, second, that it can identify anisotropic patterns of the type that are routinely missed by the rectilinear PCF. As a natural extension, we define PCFs for higher dimensions and other types of tessellations (cubic, triangular, and hexagonal) that have not been considered previously. We name these the *triangle* PCF, the *hexagon* PCF, the *cube taxicab* PCF, and the *cube uniform* PCF after the lattice set up and metric type.

Finally, we extend the concept of a PCF by introducing the *general* PCF. This PCF can be defined using any metric on any discrete domain type, with the caveat that it is more computationally expensive. We give an example of how we can use this PCF on a discrete irregular lattice (both tessellation and domain shape), where we define adjacent sites to be at unit distance from one another. We show how our PCF can identify aggregation and segregation on an irregular domain using some synthetic examples.

The paper is organized as follows. In Sec. II we discuss the successes and limitations of previous on-lattice PCFs. In Sec. III we introduce our square taxicab and square uniform PCFs. We apply our square taxicab and square uniform PCFs to some relevant examples and make comparisons with previous on-lattice PCFs from the literature in Sec. IV. In Sec. V A we define the triangle PCF, the hexagon PCF, the cube taxicab PCF, and the cube uniform PCF. We extend our PCF to more generic and possibly irregular lattices by defining the general PCF in Sec. VI. For reference, in Sec. S.4 of the Supplemental Material, we supply a table summarizing all the formulae for the normalizations of our PCFs [31]. Finally, we conclude in Sec. VII by summarising the relevance of our results and discussing potential avenues for future work.

II. EXISTING ON-LATTICE PAIR CORRELATION FUNCTIONS

In this section we provide a summary of the only two existing PCFs defined for discrete domains: the annular PCF and the rectilinear PCF. For each, we describe their strengths and limitations.

First, consider a system of agents on a two-dimensional square lattice of size $L_x \times L_y$, with lattice step Δ, and with the exclusion property that, at any given time, each lattice site can be occupied by at most one agent. If N agents occupy the domain, then the occupancy of the lattice can be represented by a matrix M:

$$M_{xy} = \begin{cases} 0 & \text{if lattice site } (x,y) \text{ is vacant,} \\ 1 & \text{if lattice site } (x,y) \text{ is occupied,} \end{cases} \tag{1}$$

where

$$N = \sum_{x=1}^{L_x} \sum_{y=1}^{L_y} M_{xy} \leqslant L_x L_y. \tag{2}$$

Let ψ^M be the set of all agent pairs in the lattice defined by matrix M, i.e.,

$$\psi^M = \big\{(\boldsymbol{a},\boldsymbol{b}) \in \mathbb{L} \times \mathbb{L} | \boldsymbol{a} = (x_a, y_a), \\ \boldsymbol{b} = (x_b, y_b),\ a \neq b,\ M_{x_a,y_a} = M_{x_b,y_b} = 1\big\}, \tag{3}$$

where $\mathbb{L} = \{1,\ldots,L_x\} \times \{1,\ldots,L_y\}$ is the set of all sites in the lattice. With agents in configuration M, let us define the subset of agent pairs separated by distance m according to some (as yet unspecified) definition of distance, denoted by d, as

$$C_d(m) = \{(\boldsymbol{a},\boldsymbol{b}) \in \psi^M | \|\boldsymbol{a}-\boldsymbol{b}\|_d = m\}, \text{ for } m \in \mathcal{D}_d, \tag{4}$$

where $\mathcal{D}_d$ is the set of possible distances under the metric d. We define the total number of pairs of agents for each value of distance $m \in \mathcal{D}_d$ as

$$c_d(m) = |C_d(m)|. \tag{5}$$

Similarly, we define the set of pairs of sites (regardless of their occupancy) which are separated by distance m according to the metric d as

$$S_d(m) = \{(\boldsymbol{a},\boldsymbol{b}) \in \mathbb{L} \times \mathbb{L} \big| \|\boldsymbol{a}-\boldsymbol{b}\|_d = m\}, \text{ for } m \in \mathcal{D}_d, \tag{6}$$

hence the total number of pairs of sites at distance m is given by

$$s_d(m) = |S_d(m)|. \tag{7}$$

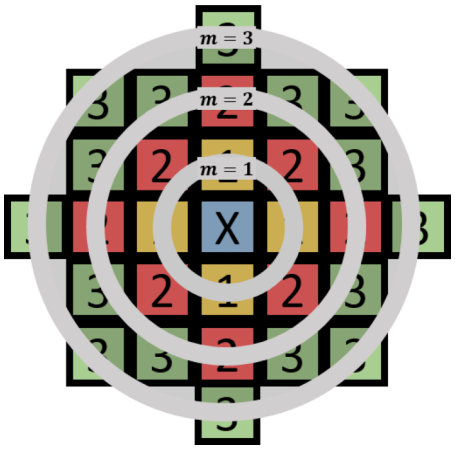


FIG. 1. Example agent pairs in the sets $C_A(m)$ with $m = 1,2,3$ and bandwidth $\delta = 1$. Concentric annuli, $(m-\delta,m]$, are superposed on top of the square lattice and the sites whose centres fall into each annulus are colored differently. Sites in yellow, red, and green (labeled 1, 2, and 3, respectively) are defined to be distance one, two, and three from the blue site, respectively, labeled X.

To produce a PCF, we aim to normalize the index $c_d(m)$ with the number of pairs we would expect at distance m if the system had no spatial correlation. That is, we consider the case in which the same number of agents are displaced uniformly at random on the lattice and compute the expected number of pairs at distance m. Let U be a random matrix such that $U_{xy} = 0$ for all sites $(x,y) \in \mathbb{L}$, apart from N sites chosen uniformly at random without replacement, for which $U_{xy} = 1$. Then we define

$$\bar{C}_d(m) = \{(\boldsymbol{a},\boldsymbol{b}) \in \psi^U \,|\, \|\boldsymbol{a}-\boldsymbol{b}\|_d = m\}, \quad \text{for} \quad m \in \mathcal{D}_d. \tag{8}$$

Hence, for each value of $m \in \mathcal{D}_d$, the PCF at distance m is defined as

$$f_d(m) := \frac{c_d(m)}{\mathbb{E}[\bar{c}_d(m)]}, \tag{9}$$

where $\bar{c}_d(m) = |\bar{C}_d(m)|$ and $\mathbb{E}$ represents the expectation operator.

A. The annular PCF

The annular PCF was originally designed for two-dimensional off-lattice systems [19] but can be extended to on-lattice systems with periodic boundary conditions (BC). For the annular PCF, the set C_A, where A denotes the annular metric, is defined as follows:

$$C_A(m) = \{(\boldsymbol{a},\boldsymbol{b}) \in \psi^M \,|\, \sqrt{(x_a-x_b)^2+(y_a-y_b)^2} \in (m-\delta,m]\}, \tag{10}$$

where $m \in \{\delta k \,|\, k \in \mathbb{N}^+\}$ and δ is a small real number which determines the bandwidth of the PCF. The schematics in Fig. 1 show a representation of some elements in the sets $C_A(1)$, $C_A(2)$, and $C_A(3)$ with $\delta = 1$.

The normalization factor is given by

$$\mathbb{E}[\bar{c}_A(m)] \approx \frac{N(N-1)(2\pi m\delta)}{L_x L_y}, \tag{11}$$

where $2\pi m\delta$ approximates the area of the mth annulus (assuming small δ) from any given agent. The annular PCF, f_A, follows from definition Eq. (9).

The normalization in Eq. (11) is a good approximation for a continuous domain with a small δ. However, when the agents are positioned on a lattice, this approach is no longer appropriate. The main issue is that the counts of agents in each annulus vary in an unpredictable manner with the distance, m, and the annular width, δ. For example, consider a square lattice with spacing Δ. The only possible distances two agents can be separated by are in the countable set

$$\mathcal{D}_A = \{\Delta\sqrt{x^2+y^2} \,|\, (x,y) \in \mathbb{N}^2 \backslash \{0,0\}\} = \{\Delta, \sqrt{2}\Delta, 2\Delta, \sqrt{5}\Delta, \dots\}. \tag{12}$$

Partitioning these distances into regularly spaced intervals, as it is required by the Euclidean distance metric, we can see that the number of agent pairs does not increase smoothly with the distance, m. Depending on the value of δ it may not even increase monotonically. However, the definition of the normalization factor Eq. (11) suggests that the expected number of pairs increases smoothly and monotonically with both m and δ. This disparity means the on-lattice annular PCF will not be properly normalized and will either be an over- or under- approximation, making results hard to interpret (see Fig. 7(b) as an example).

B. The Rectilinear PCF

In more recent work, Binder and Simpson [12,21] define the Rectilinear PCF specifically for two-dimensional, on-lattice exclusion processes with nonperiodic BC. Their definition is easily extendible to periodic BC. They define two PCFs for the two Cartesian directions. In each case the distance is defined by the number of columns (or rows) separating two agents.

Thus, the set of pairs of agents separated by integer distance $m \in \mathbb{N}^+$ are defined in the x direction and y direction respectively as

$$C_{R_x}(m) = \{(\boldsymbol{a},\boldsymbol{b}) \in \psi^M \,|\, |x_a - x_b| = m\}, \tag{13a}$$

$$C^M_{R_y}(m) = \{(\boldsymbol{a},\boldsymbol{b}) \in \psi^M \,|\, |y_a - y_b| = m\}, \tag{13b}$$

where subscripts R_x, R_y refer to the metrics defined by the Rectilinear PCFs. The schematics in Fig. 2 represent examples of sites separated by distances $m = 0$, $m = 1$ and $m = 2$ for metrics R_x and R_y.

The counts are then normalized by the expected number of pairs of agents at distance m assuming N uniformly distributed agents:

$$\mathbb{E}\big[\bar{c}_{R_x}(m)\big] = \frac{N-1}{L_x L_y - 1}\frac{N}{L_x L_y} L_y^2 (L_x - m), \tag{14a}$$

$$\mathbb{E}\big[\bar{c}_{R_y}(m)\big] = \frac{N-1}{L_x L_y - 1}\frac{N}{L_x L_y} L_x^2 (L_y - m), \tag{14b}$$

respectively. For details of the derivation of these factors, see Ref. [12]. The final rectilinear PCF is defined as the arithmetic

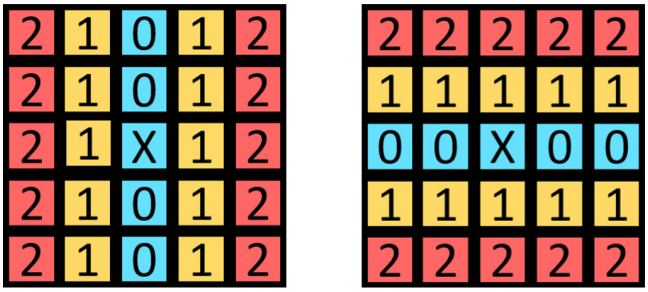


FIG. 2. Schematic of agent pairs using (a) the R_x metric, (b) the R_y metric. Sites in yellow and red (labelled 1, 2 respectively) are defined to be distance-one and -two neighbors, respectively, from the blue site labeled X.

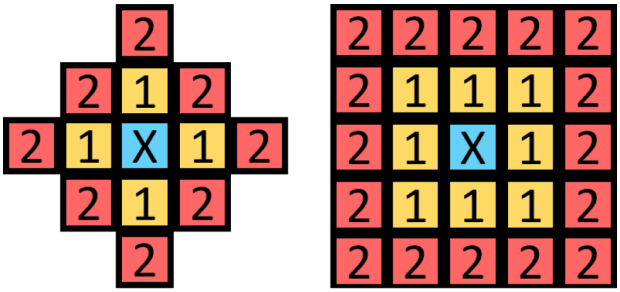


FIG. 3. Schematic of agent pairs using (a) the taxicab metric (b) the uniform metric. Sites in yellow and red, labeled with numbers 1 and 2, respectively, are defined to be distance-one and -two neighbors from the site marked in blue (labeled with X).

average of the two orthogonal PCFs, i.e.,

$$f_R^M(m) = \frac{1}{2}\big[f_{R_x}(m) + f_{R_y}(m)\big],$$

where

$$f_{R_x}^M(m) = \frac{c_{R_x}}{\mathbb{E}\big[\bar{c}_{R_x}(m)\big]},$$
$$f_{R_y}^M = \frac{c_{R_y}}{\mathbb{E}\big[\bar{c}_{R_y}(m)\big]}. \tag{15}$$

The rectilinear PCF correctly identifies spatial correlation in many examples and, unlike the annular PCF, is normalized correctly. However, one major issue that the rectilinear PCF suffers from is that, due to the inherent anisotropy of its definition, spatial structures that are biased in either Cartesian direction may be missed. For such patterns, the PCF given by the R_x and R_y metrics are approximately constant functions of distance, because the averaged row and column densities are constant along the axes, despite the fact that clustering can still be present. Examples of these spatial patterns include many biologically and chemically relevant cases, such as diagonal stripes and chessboard patterns [32,33] (see Fig. 8). We note that when the pattern structure is biased in only one Cartesian direction, the preaveraging rectilinear PCFs f_{R_x} and f_{R_y} will identify further information about the direction of the spatial pattern. Another limitation of the rectilinear PCF is that it applies only to regular square lattices and a generalisation to other forms of tessellations would be challenging.

III. THE SQUARE TAXICAB AND SQUARE UNIFORM PCFS

In this section we define two discrete PCFs for a square lattice: the square taxicab and square uniform PCF, using the taxicab and uniform metric, respectively, under both periodic and nonperiodic BC. Using the same notation as in Sec. II, we define the subsets of agent pairs separated by distance m under nonperiodic BC as

$$C_1^n(m) = \{(\boldsymbol{a},\boldsymbol{b}) \in \psi^M \,|\, \|\boldsymbol{a}-\boldsymbol{b}\|_1 = m\}, \quad m \in \mathcal{D}_1^n, \tag{16a}$$
$$C_\infty^n(m) = \{(\boldsymbol{a},\boldsymbol{b}) \in \psi^M \,|\, \|\boldsymbol{a}-\boldsymbol{b}\|_\infty = m\}, \quad m \in \mathcal{D}_\infty^n, \tag{16b}$$

for the taxicab and uniform metric, respectively. Here, $\mathcal{D}_1^n = \mathcal{D}_\infty^n = \{1,2,\ldots \min\{L_x,L_y\}-1\}$ and the superscript n refers to the fact we are considering *nonperiodic* BC. Using the definitions of the uniform and taxicab metrics, we can express these sets as

$$C_1^n(m) = \{(\boldsymbol{a},\boldsymbol{b}) \in \psi^M \,|\, |x_a - x_b| + |y_a - y_b| = m\}, \tag{17a}$$
$$C_\infty^n(m) = \{(\boldsymbol{a},\boldsymbol{b}) \in \psi^M \,|\, \max\{|x_a - x_b|, |y_a - y_b|\} = m\}. \tag{17b}$$

Similarly, we define the subsets of agent pairs separated by distance m under periodic BC as

$$C_1^p(m) = \{(\boldsymbol{a},\boldsymbol{b}) \in \psi^M \,|\, \min\{|x_a - x_b|, L_x - |x_a - x_b|\} + \min\{(|y_a - y_b|, L_y - |y_a - y_b|\} = m\}, \quad m \in \mathcal{D}_1^p, \tag{18a}$$
$$C_\infty^p(m) = \{(\boldsymbol{a},\boldsymbol{b}) \in \psi^M \,|\, \max\{\min\{|x_a - x_b|, L_x - |x_a - x_b|\}, + \min\{(|y_a - y_b|, L_y - |y_a - y_b|\}\} = m\}, \quad m \in \mathcal{D}_\infty^p, \tag{18b}$$

where $\mathcal{D}_1^p = \mathcal{D}_\infty^p = \{1,2,\ldots \min\{\lfloor \frac{L_x}{2} \rfloor, \lfloor \frac{L_y}{2} \rfloor\}\}$. The corresponding definitions of S_1^n, S_∞^n, S_1^p, and S_∞^p can be obtained similarly by using Eq. (6). Here the superscript p refers to the fact we are considering *periodic* BC. Notice that we restrict the largest m to be $\min\{\lfloor \frac{L_x}{2} \rfloor, \lfloor \frac{L_y}{2} \rfloor\}$ to simplify the computation of the normalization factor (see Sec. III A). However, with some work this restriction could be relaxed. The schematics in Fig. 3 represent examples of sites separated by distance $m = 1$ and $m = 2$ using the taxicab (a) and uniform (b) metrics.

For the normalization factors, Binder and Simpson [12] use

$$\mathbb{E}[\bar{c}_d(m)] = \left(\frac{N}{L_x L_y}\right)\left(\frac{N-1}{L_x L_y - 1}\right) s_d(m), \tag{19}$$

where $s_d(m)$ is defined as in Eq. (7) and d refers to the metric used. In other words, the expected number of pairs of agents at distance m on a lattice with N uniformly distributed agents can

be written as the probability that two different sites at distance m are simultaneously occupied, multiplied by the total number of pairs of sites at distance m in the domain.

To complete the definitions of our square lattice PCFs we need to provide an expression for s_1 and s_∞. We address this in the next two sections, distinguishing between the cases of periodic and nonperiodic BC.

A. Normalization of the square taxicab and square uniform PCF under periodic boundary conditions

As the derivation for the normalization is simple under periodic BC and more complicated under nonperiodic BC, we first consider the system with periodic BC and determine s_1^p, s_∞^p, where p denotes periodic BC. Let us define the number of sites separated by a distance m from any given reference site on a lattice as $t_1(m)$, $t_\infty(m)$ under the taxicab and uniform metric, respectively. These read

$$t_1(m) = 4m, \tag{20a}$$

$$t_\infty(m) = 8m. \tag{20b}$$

The proofs of Eqs. (20) are omitted, but they can be obtained easily by induction on m. Examples for $m = 1,2$ can be seen in Fig. 3. Notice that for $m \leqslant \min\{\lfloor \frac{L_x}{2} \rfloor, \lfloor \frac{L_y}{2} \rfloor\}$, given any site on the lattice, the number of sites at distance m from this reference site in the case of periodic BC is exactly $t(m)$. Consider the lattice of size $L_x \times L_y$ with $L_x, L_y > 2$. If we multiply the total number of lattice sites by $t(m)$, we count each pair of sites separated by distance m exactly twice. Hence, we conclude that

$$s^p_{d=1,\infty}(m) = \frac{t(m)L_xL_y}{2}, \tag{21}$$

using the taxicab and uniform metrics. Substituting values for $t_1(m)$ and $t_\infty(m)$ from Eqs. (20) we deduce that

$$s_1^p(m) = 2mL_xL_y, \tag{22a}$$

$$s_\infty^p(m) = 4mL_xL_y. \tag{22b}$$

Therefore, by substituting Eqs. (22) into Eq. (19), the normalization factors under periodic BC are

$$\mathbb{E}\left[\bar{c}_1^p(m)\right] = \frac{2mN(N-1)}{L_xL_y - 1}, \tag{23a}$$

$$\mathbb{E}\left[\bar{c}_\infty^p(m)\right] = \frac{4mN(N-1)}{L_xL_y - 1}. \tag{23b}$$

B. Normalization of the square taxicab and square uniform PCF under nonperiodic BC

In this section we derive expressions for $s_1^n(m)$ and $s_\infty^n(m)$, where n denotes nonperiodic BC. Notice that, for all $m \in \mathcal{D}$, we have that $s^p(m) > s^n(m)$ since $s^p(m)$ includes pairs that cross the domain boundary, whereas $s^n(m)$ does not. Therefore, to find a formula for $s^n(m)$, it is enough to determine a formula for the remainders defined by

$$r_1(m) = s_1^p(m) - s_1^n(m), \tag{24a}$$

$$r_\infty(m) = s_\infty^p(m) - s_\infty^n(m). \tag{24b}$$

The remainders count the number of pairs of sites that cross a boundary under the periodic BC. For simplicity, throughout this section, we will only derive the normalization for the taxicab metric; however, the derivation for the uniform metric is similar and can be found in Supplemental Material Sec. S.1 for reference [31]. Let us define the set of pairs of sites separated by distance $m \in \mathcal{D}_1^n$ that cross the x boundary (horizontal axis) or y boundary (vertical axis), respectively, as

$$P_1^x(m) = \left\{(\boldsymbol{a},\boldsymbol{b}) \in S_1^p(m) \middle| |y_a - y_b| > L_y - |y_a - y_b|\right\}, \tag{25a}$$

$$P_1^y(m) = \left\{(\boldsymbol{a},\boldsymbol{b}) \in S_1^p(m) \middle| |x_a - x_b| > L_x - |x_a - x_b|\right\}. \tag{25b}$$

Of these pairs, let us consider those pairs that are at distance $k \in \{1,\ldots m\}$ rows or columns from each other, respectively. We define these subsets as

$$P_1^x(m,k) = \left\{(\boldsymbol{a},\boldsymbol{b}) \in P_1^x(m) \middle| L_y - |y_a - y_b| = k\right\}, \tag{26a}$$

$$P_1^y(m,k) = \left\{(\boldsymbol{a},\boldsymbol{b}) \in P_1^y(m) \middle| L_x - |x_a - x_b| = k\right\}. \tag{26b}$$

Notice that $P_1^x(m) = \bigcup_{k=1}^m P_1^x(m,k)$ and $P_1^y(m) = \bigcup_{k=1}^m P_1^y(m,k)$. Figures 4(a) and 4(b) give visualizations of pairs of sites within $P_1^x(m,m)$. Figure 4(c) gives examples of distances between pairs of sites in $P_1^x(m,k)$, for $k = 1,\ldots,m$. By definition Eqs. (16a) and (18a) we have that

$$S_1^p(m) \backslash S_1^n(m) = P_1^x(m) \cup P_1^y(m). \tag{27}$$

Hence, by combining Eqs. (24a) and (7), we obtain

$$\begin{aligned} r_1(m) &= \left|P_1^x(m) \cup P_1^y(m)\right| \\ &= \sum_{k=1}^m \left|P_1^x(m,k)\right| + \sum_{k=1}^m \left|P_1^y(m,k)\right| - \left|P_1^x(m) \cap P_1^y(m)\right|. \end{aligned} \tag{28}$$

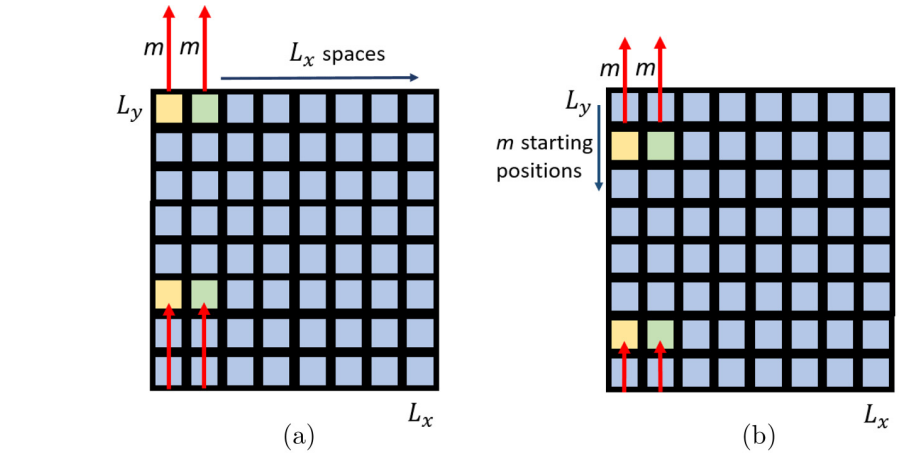
To conclude the computation we derive an expression for the two sums in Eq. (28) and the corresponding equation for the size of the intersection. By counting the contribution of each type of pair (see Fig. 4 for a visualization), one can write down the following expressions for the two sums in Eq. (28):

$$\sum_{k=1}^m \left|P_1^x(m,k)\right| = 2[L_x + 2L_x + \cdots + L_x(m-1)] + L_xm, \tag{29a}$$

$$\sum_{k=1}^m \left|P_1^y(m,k)\right| = 2[L_y + 2L_y + \cdots + L_y(m-1)] + L_ym\ . \tag{29b}$$

Hence,

$$\begin{aligned} &\sum_{k=1}^m \left|P_1^x(m,k)\right| + \sum_{k=1}^m \left|P_1^y(m,k)\right| \\ &\quad = 2[L_x + 2L_x + \cdots + L_x(m-1)] + L_xm \\ &\qquad + 2[L_y + 2L_y + \cdots + L_y(m-1)] + L_ym \end{aligned}$$



k	1	2	...	$m-1$	m
Illustration of pair types in $P_1^x(m,k)$	m-1 1 m-1	m-2 2 m-2		1 m-1 1	m
$\lvert P_1^x(m,k)\rvert$	$2(L_x)$	$2(2L_x)$	...	$2((m-1)L_x)$	mL_x

(c)

FIG. 4. A visualization of the pairs of sites in $P_1^x(m)$. Panels (a) and (b) show two different site pairs in $P_1^x(m,m)$. For each of the L_x columns, each site in the rows $\{L_y-m+1, L_y-m+2, \ldots, L_y\}$ has a single corresponding site at distance m separated by m rows and 0 columns and reached by crossing the horizontal boundary. Therefore $|P_1^x(m,m)| = mL_x$. Panel (c) shows all the different types of pairs in $P_1^x(m,k)$ for $k = 1, \ldots, m$ with the corresponding value of $|P_1^x(m,k)|$.

$$
\begin{aligned}
&= (L_x + L_y)\left(m + 2\sum_{i=1}^{m-1} i\right) \\
&= (L_x + L_y)\left[m + 2\frac{(m-1)m}{2}\right] \\
&= (L_x + L_y)m^2. \qquad (30)
\end{aligned}
$$

We now focus on deriving an expression for the size of intersection, $|P_1^x(m) \cap P_1^y(m)|$, in Eq. (28). The set $P_1^x(m) \cap P_1^y(m)$ consists of pairs of sites separated by distance m that cross both the x and y boundaries simultaneously.

There are two regions of the domain where site pairs cross two boundaries. These are any two consecutive corners of the four corners of the domain. Examples of these regions and site pairs within these regions are visualized in Fig. 5. In Fig. 6 we give an illustrative example in which we count the number of these pairs for $m = 5$. All sites inside the boundaries of the domain colored in orange, purple, green, and yellow are distance $m = 5$ from other sites of the same color outside the boundaries of the domain. Notice that the yellow site in the corner at (L_x, L_y) is distance five from a total of four sites, reached by crossing the x and y boundaries, denoted by a 4 in the site. Similarly, the two green sites at $(L_x - 1, L_y)$ and $(L_x, L_y - 1)$ are distance m from three sites, reached by crossing the x and y boundary, denoted by a 3 in the two sites. $|P_1^x(5) \cap P_1^y(5)|$ is the sum of all the numbers in the colored sites multiplied by two to account for the second corner region. Extrapolating, for any value of m, the number of pairs of sites

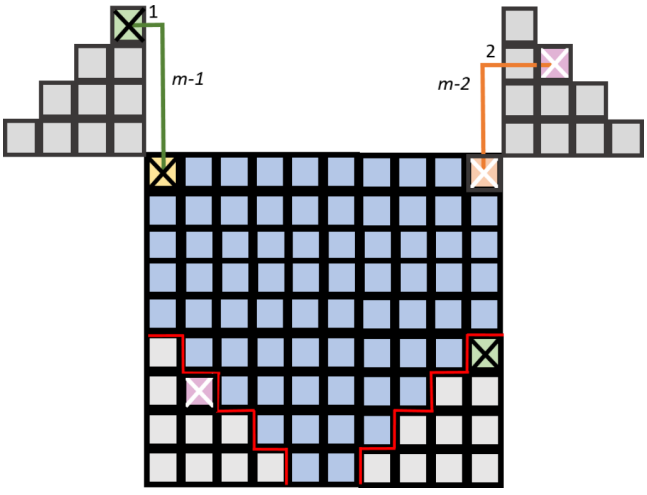


FIG. 5. Examples of pairs of sites separated by distance $m = 5$ that cross both the x and y boundaries, i.e., pairs of sites in $P_1^x(m) \cap P_1^y(m)$. The gray sites outside the domain correspond to the gray sites inside the domain in the diametrically opposite corner. As illustrations of site pairs at a distance m which cross both boundaries, the orange site containing a white cross is distance m from the pink site containing a white cross. Similarly, the yellow site containing a black cross is distance m from the green site containing a black cross.

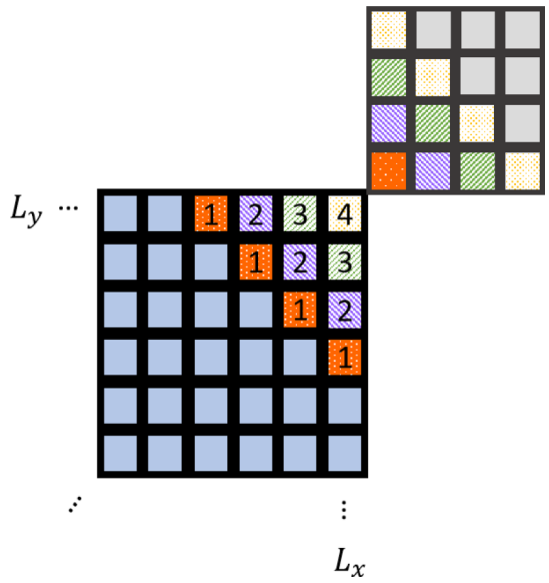


FIG. 6. Examples of pairs in $P_1^x(5) \cap P_1^y(5)$ on a zoomed-in corner of a larger domain. Sites with a given pattern (not plain) are distance $m = 5$ away from sites with the same pattern that can be reached by crossing the x and y boundaries. The number in each given site corresponds to the number of sites at distance $m = 5$, reached by crossing the x and y boundaries.

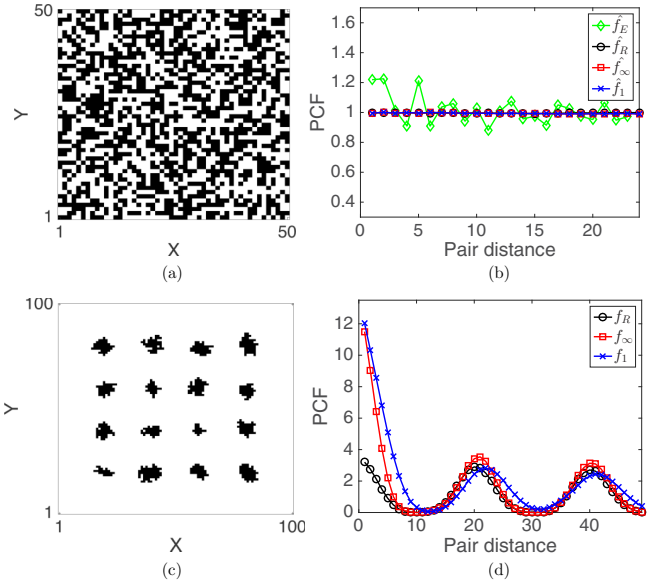


FIG. 7. Examples of spatial structure analysis. In panels (a) and (b) a case with no spatial correlation is considered. Panel (a) is an example visualization of an occupancy matrix uniformly populated with density 0.5. Occupied sites are colored in black and are white otherwise. Panel (b) shows the four PCFs each averaged over 50 uniformly populated matrices. Panels (c) and (d) refer to a discrete simulation of the agent-based model described in the text, at time $t = 10$. Panel (c) is a visualization of the occupancy matrix and in panel (d) our PCFs are compared with the Rectilinear PCF.

that cross the two boundaries is exactly

$$\left|P_1^x(m) \cap P_1^y(m)\right| = 2[(m-1) + 2(m-2) + 3(m-3) + \cdots + m - 1] = 2\sum_{i=1}^{m-1}(m-i)i = \frac{m^3 - m}{3}. \quad (31)$$

By substituting Eqs. (31) and (30) into Eq. (28) we gain an expression for the remainder $r_1(m)$. By rearranging Eq. (24a) we determine $s_1^n(m)$, which we then substitute into Eq. (19) to obtain the exact expression for the normalization in the nonperiodic case. This is given by

$$\mathbb{E}\left[\bar{c}_1^n(m)\right] = \left(\frac{N}{L_x L_y}\right)\left(\frac{N-1}{L_x L_y - 1}\right)\left[2mL_xL_y - (L_x + L_y)m^2 + \frac{m^3 - m}{3}\right]. \quad (32)$$

A similar approach can be used to obtain the normalization factor for the uniform metric under nonperiodic BC:

$$\mathbb{E}\left[\bar{c}_\infty^n(m)\right] = \left(\frac{N}{L_x L_y}\right)\left(\frac{N-1}{L_x L_y - 1}\right)[4mL_xL_y - 3(L_x + L_y)m^2 + 2m^3]. \quad (33)$$

For more details on the derivation of Eq. (33), see Supplemental Material Sec. S.1.

IV. RESULTS

In this section we use the square taxicab and square uniform PCFs defined in Sec. III to analyze the spatial correlation in some examples. We compare our results with previously suggested on-lattice PCFs.

We start by computing the PCFs for a system without any spatial correlation. We consider 50 independent occupancy matrices, U_i, $i = 1, \ldots, 50$, populated uniformly at random with density 0.5 [see Fig. 7(a) for example]. For each realisation, U_i, we compute the corresponding PCF, $f_d^{U_i}$, and then we average the results over the 50 realisations which we denote $\hat{f}_d$. If the normalization is correct, $\hat{f}_d(m)$ should return the value unity for every pair distance, m, meaning that no spatial correlation is found. In Fig. 7(b) all four aforementioned averaged PCFs are plotted: $\hat{f}_A$, $\hat{f}_R$, $\hat{f}_1$, and $\hat{f}_\infty$. The results show that both the averaged square uniform PCF, $\hat{f}_1$, and square taxicab PCF, $\hat{f}_\infty$, correctly predict that there is no spatial correlation. The averaged rectilinear PCF, $\hat{f}_R$, also correctly predicts no spatial correlation. However, the averaged annular PCF, $\hat{f}_A$, has clear peaks, suggesting, incorrectly, the presence of spatial correlation. Since the results are averaged over multiple repeats, such a discrepancy can not be attributed to stochasticity, but is due to incorrect normalization as explained in Sec. II. Note that the annular PCF can still correctly identify spatial correlation in many examples; however, the incorrect normalization often makes the results hard to interpret. This is because the annular PCF makes it difficult to distinguish between genuine correlation and systematic error. For this reason, for the rest of this section, we omit the results of the annular PCF and continue to compare between our PCFs and the rectilinear PCF using nonperiodic BC.

Next we consider examples of strong spatial correlation. Figure 7(c) shows an example of aggregation driven by a

proliferation mechanism. The occupancy matrix is obtained by simulating an on-lattice agent-based model with periodic BC as described in Binder and Simpson [12], which we summarize as follows. The model is initialized with 16 agents, located at coordinates given by $\{(x,y)\,|\,x,y \in \{20,40,60,80\}\}$ on a regular square lattice with $L_x = 100$, $L_y = 100$. Time is discretized with a time step $\tau = 1$ and the number of agents at time t is denoted by $n(t)$. At each time step the configuration at time $t+\tau$ is obtained from the configuration at time t, by repeating the following steps $n(t)$ times. (1) An agent is chosen uniformly at random from the $n(t)$ agents present at the end of the previous time-step; (2) one of its four von Neumann neighbors is selected at random with equal probability; (3) if the selected site is empty, then a new agent is placed in this site and $n(t+\tau) = n(t)+1$; otherwise, the configuration is left unchanged.

Figure 7(c) shows a single realisation after 10 time steps and Fig. 7(d) shows the corresponding PCFs: f_R, f_1, and f_∞. The results indicate that all of the PCFs correctly identify aggregation. However, the quantitative information about aggregate sizes at different length scales provide by each PCF varies. For example, we see that all PCFs in Fig. 7(d) exhibit three peaks; at $m = 1$, $m \approx 20$, and $m \approx 40$. The different peaks and troughs of the PCF profiles have different qualitative meanings related to the correlation type. Due to the local approach of the square taxicab PCF and square uniform PCF, the first peak at $m = 1$ is three times higher than the peaks at larger values of distance. These differences in amplitude highlight the different peak origins. Specifically, the first and highest peak distinguishes the individual cluster aggregate and the later peaks indicate correlation between different clusters. In contrast, all three peaks in the rectilinear PCF are the same amplitude. Note that, in the case of aggregation, the average diameter of the aggregate corresponds to the first value of distance which achieves the minimum of the PCF. The rectilinear, square uniform, and square taxicab PCFs estimate the aggregate diameter to be 9, 9 and 11, respectively. Importantly, the PCFs capture the fact that this diameter depends on the metric used. In particular the distance between two sites measured using the uniform metric is always less than or equal to the taxicab distance. This phenomenon is seen more clearly in later examples.

We now consider a series of examples with spatial correlation constructed artificially to compare and evaluate the different PCFs. In Fig. 8 we compare our square uniform and square taxicab PCF with the rectilinear PCF for three different patterns with strong spatial correlation. All three examples (Fig. 8(a), diagonal stripes; Fig. 8(c), chessboard pattern; and Fig. 8(e), concentric circles) are chosen so that the column- and row-averaged densities are constant and hence the spatial structure is not recognised by the rectilinear PCF, as shown in Figs. 8(b), 8(d), and 8(f). This is in contrast to the approach of our PCFs (both square uniform and square taxicab), which successfully recognize the spatial structure in all three examples. In addition, these examples uncover other interesting differences between the taxicab and the uniform approaches. Consider the PCF for the case of diagonal stripes and the chessboard pattern [Figs. 8(b) and 8(d)]. Here the square uniform PCF quickly converges to unity (no spatial correlation) for large distance m, while in both cases, the square taxicab PCF still shows a strong oscillatory behavior for large distance m, suggesting spatial correlation. To give an intuitive

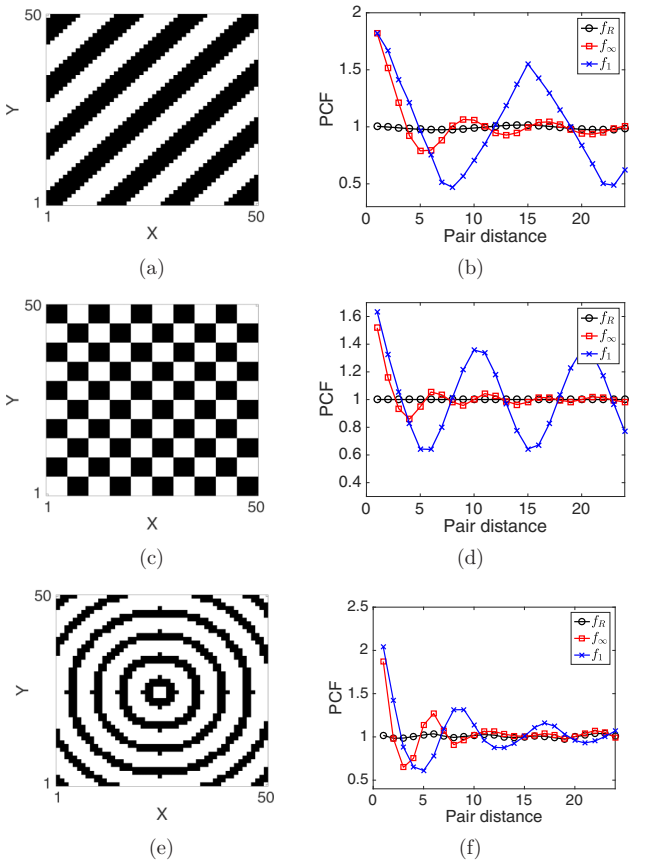


FIG. 8. Examples of pattern analysis. Panels (a), (c), and (e) visualize three constructed spatial patterns. Panels (b), (d), and (f) displays the corresponding square taxicab, square uniform and rectilinear PCF using the occupancy matrices for (a), (c), and (e), respectively.

explanation of this phenomenon, let us consider the shapes of balls of size m centered at a given site $\boldsymbol{a}$ under the two metrics. These balls are defined as $B_m(\boldsymbol{a}) = \{\boldsymbol{b} \in \mathbb{L} \,|\, \|\boldsymbol{a}-\boldsymbol{b}\|_d \leqslant m\}$ with $d = 1,\infty$, respectively (see Fig. 3). The ball corresponding to the uniform metric [Fig. 3(b)] has a square shape with the sides aligned with the directions of the axis. This implies that when distance m becomes close to either L_x or L_y in size, the ball corresponding to the uniform metric of distance m contains most of the sites in the corresponding row or column at distance m. For large m, therefore, the uniform metric begins to work in a similar way to the Rectilinear PCF and thus fails to recognize anisotropic patterns biased in the Cartesian directions. The ball of the taxicab metric [see Fig. 3(a)], however, has a diamond shape. Consequently, the long-distance correlations appear clear even for patterns in which both the average column and row densities are constant, as in Fig. 8.

The examples in Fig. 8 were constructed specifically to underline the main differences between the three PCFs. Nevertheless, similar patterns also arise in many biologically and mathematically relevant applications [32–34]. We conclude this section by comparing the three PCF approaches applied to some real-world examples taken from the literature. In Fig. 9 we analyze three images representing examples of Turing patterns. A corresponding occupancy matrix is obtained by representing each pixel of the image as a value in a matrix

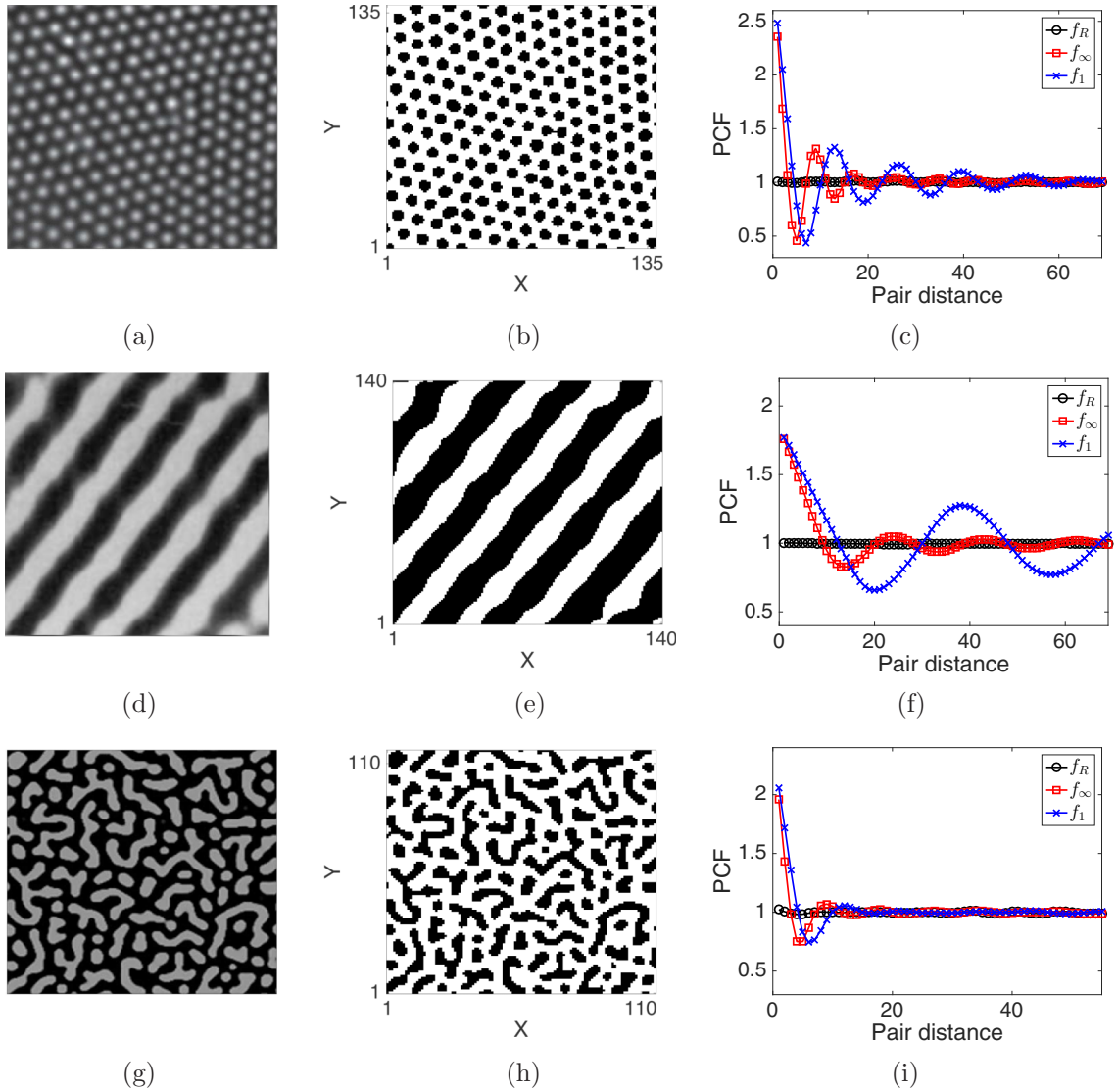


FIG. 9. Spatial analysis of Turing patterns. Panels (a), (d), and (g) show original images representing the results of a reaction-diffusion mechanism between two chemical substances, (a) is reprinted from Ref. [35], (d) from Ref. [32], and (g) from Ref. [36]. Panels (b), (e), and (h) visualize the occupancy matrices corresponding to the original images (described in text). In panels (c), (f), and (i) we compare the square taxicab, square uniform, and rectilinear PCFs for each of the examples.

which is 1 (i.e., occupied) if the three values of the RGB colorization of the pixel are above a certain threshold (80) and 0 otherwise. In all cases the column and row densities are almost constant, hence the spatial structure again remains largely undetected by the rectilinear PCF, while our square uniform and square taxicab PCF correctly identify the patterns. As already observed in the previous examples, we note that the estimated diameter of the aggregate, wavelength, and amplitude of the oscillations differ according to the metric used.

V. THE TRIANGLE, HEXAGON, AND CUBE PCFS

Despite the square lattice being the most popular set up for spatially discrete models [13,37–40], in some situations other types of tessellation, either regular or irregular, can be more suitable [13,17,41].

In the following subsections we extend our definition of the PCFs in Sec. III to more general types of tessellations. We define the triangle, hexagon, cube uniform, and cube taxicab PCFs under nonperiodic and periodic BC for triangular, hexagonal, and cuboidal tessellations, respectively. The following subsections represent qualitative discussions of the different cases. We refer the reader to the Supplemental Material Secs. S.2 and S.3 for the full details of the derivation of the PCF formulas [31].

A. Triangle and hexagon PCF

First, we define triangularly and hexagonally tessellated domains of size $L_x \times L_y$. These comprise an array of L_y rows of L_x regular triangles or hexagons, respectively. Examples for which $L_x = 6$ and $L_y = 3$ for each of the two cases are given in Figs. 10(a) and 10(b), respectively. Notice that, for a

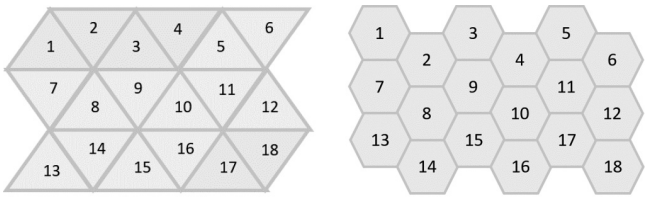


FIG. 10. Example domains for (a) triangular and (b) hexagonal tessellations in which $L_x = 6$ and $L_y = 3$.

periodic BC to be meaningful in these domain definitions, L_x must be even. Therefore, we enforce this as a condition in what follows.

In the context of triangular and hexagonal tessellations we focus our attention on the taxicab metric, both for simplicity and as the most natural metric on this domain type. Using the taxicab metric, the number of sites of distance m from any given reference site is given by

$$t_{\text{tri}}(m) = 3m, \tag{34a}$$

$$t_{\text{hex}}(m) = 6m. \tag{34b}$$

The proofs of Eqs. (34) are omitted, but they can be obtained easily by induction on m. Examples for $m = 1,2,3$ are visualized in Fig. 11. Using the same reasoning as in Sec. III under periodic BC:

$$s^p_{\text{tri}}(m) = 3m\frac{L_x L_y}{2}, \tag{35a}$$

$$s^p_{\text{hex}}(m) = 3m L_x L_y. \tag{35b}$$

Substituting Eqs. (35) into Eq. (19) we obtain the normalizations for the triangle and hexagon PCF, respectively, under periodic BCs, namely,

$$\mathbb{E}\left[\bar{c}^p_{\text{tri}}(m)\right] = \frac{3mN(N-1)}{2(L_x L_y - 1)}, \tag{36a}$$

$$\mathbb{E}\left[\bar{c}^p_{\text{hex}}(m)\right] = \frac{3mN(N-1)}{L_x L_y - 1}. \tag{36b}$$

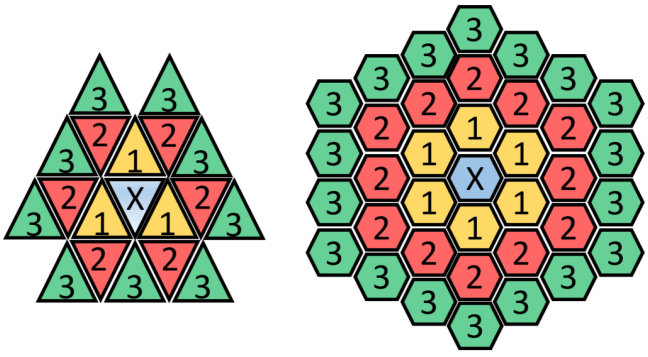


FIG. 11. Schematic of agent pairs with the (a) triangular tessellation and (b) hexagonal tessellation using the taxicab metric. Sites in yellow, red, and green, labeled 1, 2, and 3, respectively, are distance-one, -two, and -three neighbors from the blue site (labeled with X), respectively.

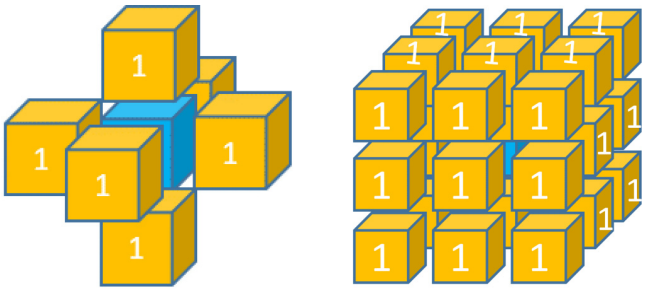


FIG. 12. Schematic of agent pairs using (a) the taxicab metric and (b) the uniform metric. Sites in yellow are defined to be distance-one neighbors from the site marked in blue.

From these expressions one can obtain the formulas of f_{tri} and f_{hex} under periodic BC by using the definition Eq. (9). The normalizations in the case of nonperiodic BC are given in the Supplemental Material Sec. S.2 [31].

B. Uniform cube and taxicab cube PCF

We define a three-dimensional $L_x \times L_y \times L_z$ cuboidal lattice with unit spacing. Using the taxicab and uniform metric, respectively, the number of sites of distance m from any given reference site is given by

$$t_{\text{cube}_1}(m) = 2(2m^2 + 1), \tag{37a}$$

$$t_{\text{cube}_\infty}(m) = 2(12m^2 + 1). \tag{37b}$$

The proofs of Eqs. (37) are omitted, but they can be obtained easily by induction on m. Examples of agent pairs for $m = 1$ are given in Figs. 12(a) and 12(b) for the taxicab and uniform metrics, respectively. Using the same reasoning as in Sec. III, under periodic BCs, the normalizations for taxicab and uniform cube PCFs, respectively, are as follows:

$$s^p_{\text{cube}_1}(m) = (2m^2 + 1)L_x L_y L_z, \tag{38a}$$

$$s^p_{\text{cube}_\infty}(m) = (12m^2 + 1)L_x L_y L_z. \tag{38b}$$

For simplicity we refer the reader to Sec. S.3 of the Supplemental Material for the normalization factors for the cases with nonperiodic BC [31].

VI. THE GENERAL PCF

In this section we provide a comprehensive method for generating a PCF for any tessellation type, BC, and metric but with the caveat of having a high computational cost.

This PCF is a valuable tool for irregular domain shapes and partitions although it can be used for any tessellation of any domain.

First, we consider a two-dimensional domain partitioned into Z regions (or sites) with arbitrary shapes and sizes, each labeled with a number from 1 to Z. Figure 13(a) shows an example of an irregularly shaped domain partitioned in $Z = 17$ regions. Given the domain, we choose a suitable metric. For the irregular lattice, which we consider in the following example, we consider the taxicab metric. This means that we define the

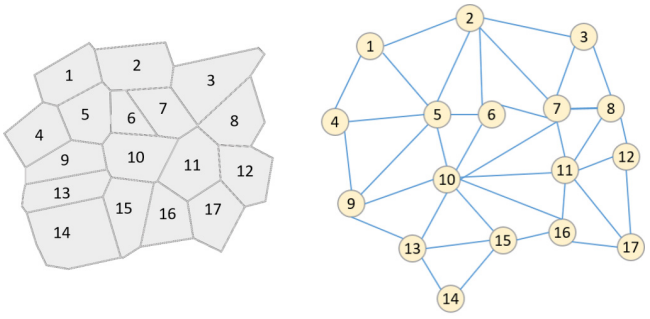


FIG. 13. An example of an irregular domain partition with its corresponding connectivity graph under the taxicab metric. Panel (a) shows A size 17 irregular lattice domain. Panel (b) shows corresponding connectivity graph for the tessellation in (a) under nonperiodic BC using the taxicab metric for distances.

distance, m, between two sites to be the minimum number of sites visited when starting at one site and moving consecutively through adjacent sites to the other. For example, in Fig. 13(a) the sites 4 and 7 are at distance three. Similarly, adjacent sites are defined to be at distance one.

Having chosen and defined a suitable metric, we may now represent the connections between lattice sites as an undirected connectivity graph $G(V,E)$, where each vertex represents a lattice site and each edge connects vertices whose corresponding sites are distance-one neighbors. Using the taxicab metric, edges connect vertices whose corresponding sites are adjacent. Figure 13(b) shows an example of such an association [applied to the irregular lattice in Fig. 13(a)] using the taxicab metric.

The corresponding adjacency matrix of graph G is a $Z \times Z$ matrix defined as follows:

$$A^G_{i,j} = \begin{cases} 1 & \text{for } (i,j) \in E, \\ 0 & \text{for } (i,j) \notin E. \end{cases} \tag{39}$$

We use properties of the adjacency matrix to determine the number of sites at a given distance. In particular, we can compute $(A^G)^m$ whose entries $(A^G)^m_{i,j}$ are the number of walks of length m from vertex i to vertex j. To compute the minimum walk between two sites (and hence the distance between them) we produce the *distance matrix* D^G. This is a $Z \times Z$ matrix defined as

$$D^G_{i,j} = \begin{cases} \min\left\{m \in \mathbb{N}^+ \middle| (A^G)^m_{i,j} \neq 0\right\} & \text{for } i \neq j, \\ 0 & \text{for } i = j. \end{cases} \tag{40}$$

Notice that, each entry, $D^G_{i,j}$, denotes the distance between the vertices i and j on graph G and hence on the original lattice.

Given the distance matrix of the domain, D^G, and the set of the occupied sites $M \subseteq V$, the PCF of the system can be computed as follow. The number of pairs of agents at distance m for a general metric d is given by

$$c_d(m) = \tfrac{1}{2}\left|\left\{(i,j) \in M \times M \middle| D^G_{i,j} = m\right\}\right|. \tag{41}$$

Similarly, we can express the number of pairs of sites at distance m as

$$s_d(m) = \tfrac{1}{2}\left|\left\{(i,j) \in V \times V \middle| D^G_{i,j} = m\right\}\right|. \tag{42}$$

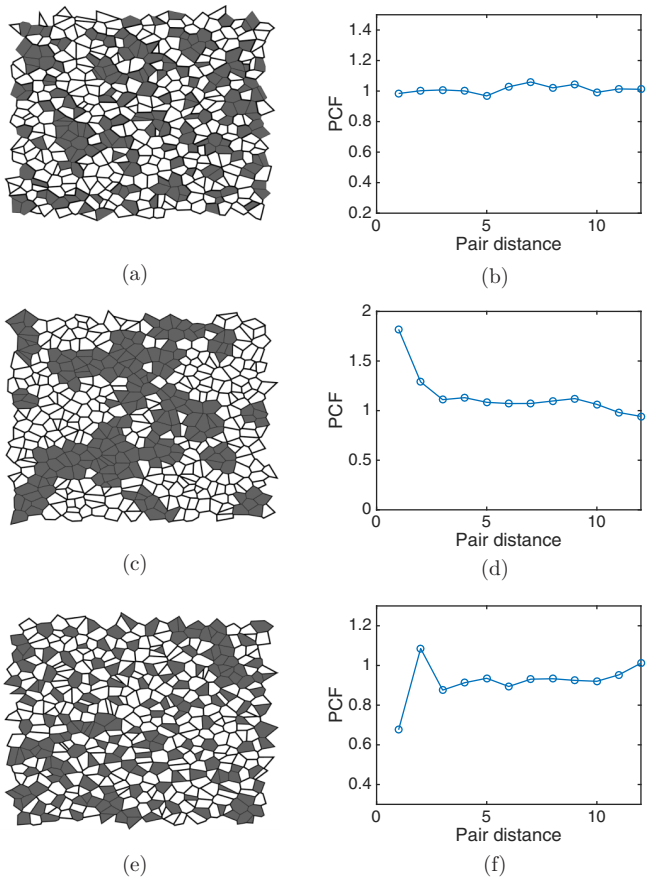


FIG. 14. Examples of spatial correlation analysis on an irregular domain. Panels (a), (c), and (e) show three examples of irregular lattices populated with density 0.4 with agents (gray sites). In panel (a), agents are displaced uniformly at random (no spatial correlation). In panel (c), agents are in a strong form of aggregation, while in panel (e) agents are displaced in a segregate manner. Panels (b), (d), and (f) are the corresponding general PCF evaluations for panels (a), (c), and (e), respectively.

To compute the normalization factor, denote the total number of agents as $N = |M|$, and hence using the same argument as in Sec. III we can write

$$\mathbb{E}[\bar{c}_d(m)] = \left(\frac{N}{Z}\right)\left(\frac{N-1}{Z-1}\right)s_d(m). \tag{43}$$

The general PCF is then defined by combining Eqs. (42) and (43) in Eq. (9). Notice that the computation of the normalization for the general PCF can be computationally expensive. This is because the computation of D^G involves calculating powers of matrices of size $Z \times Z$, where Z is often large. In particular, the cost of computing the normalization of the general PCF is $O(Z^3 m_{\max})$ in which $m_{\max}$ is the maximum value m for which the PCF is computed. For this reason, the general PCF is better reserved for cases in which the expression for the normalization factor cannot be computed analytically, unlike in Secs. III, V A, and V B, although it can, of course, be used even if an analytical formula is available.

In Fig. 14 we apply the general PCF to three examples of agent-based systems on an irregular lattice. In all three examples, the irregular tessellation is the Voronoi partition

based on a set of randomly distributed points. The randomized points are obtained by starting with a square lattice, with lattice size Δ, and perturbing the coordinates of each point (x_i, y_i) to $(x_i + \delta_i^x, y + \delta_i^y)$ with each δ_i^x, δ_i^y chosen uniformly at random in the interval $[-\frac{\Delta}{2}, \frac{\Delta}{2}]$.

In the first example, Fig. 14(a), N lattice sites selected uniformly at random to be occupied (gray sites). By eye, several larger clusters of occupied and unoccupied lattice sites are evident, indicating that there may be spatial correlation. Figure 14(b) shows the corresponding general PCF. The values are close to unity, correctly identifying that there is, indeed, no spatial correlation in the system. This highlights the importance of accurate quantitative methods for determining spatial correlation rather than a reliance on ad hoc judgements.

In the other two cases we test our PCF with examples of strong spatial correlation. In Fig. 14(c) we consider a system in an aggregated state. To generate such a configuration, we start with an empty domain, and select an empty site uniformly at random. We then place an agent in this site and all of its neighboring sites (if they are not already occupied). We repeat this process until we reach density 0.4. The process leads to a strong form of aggregation which the general PCF, in Fig. 14(d), correctly identifies. Finally, in Fig. 14(e) we consider a system in a segregated state. To generate such a configuration, we start with a fully populated domain, and repeatedly select at random an occupied site. We remove all agents occupying adjacent sites but leave the initially selected site occupied. The process ends when density 0.4 is reached. This mechanism generates a system which is unlikely to have adjacent sites occupied and more likely to have agents displaced at distance two. The corresponding PCF, shown in Fig. 14(f), correctly captures both of these features: the PCF has value 0.68 at pair distance one, which implies negative correlation at the shortest distance, and value 1.1 at pair distance two, highlighting the positive correlation at slightly larger distances.

Note that, since the normalization in the general PCF can give an exact value for the number of sites at certain pair distances, we used this method to check the normalization factors for all of our previously proposed PCFs. In all cases the results confirmed the analytical expressions given in Secs. III, V A, and V B and in the Supplemental Material [31].

VII. CONCLUSIONS

In this paper we have developed a set of tools to study spatial correlation on discrete domains. We derived two discrete pair-correlation functions for an exclusion process on a square lattice: the square uniform and square taxicab PCF. We applied our PCFs to patterns observed in nature and to computational simulations and showed that our PCF can not only distinguish and quantify different types of correlation but also that it improves upon previous on-lattice PCFs. For example, we showed that our PCF was normalized correctly, unlike the annular PCF, and was able to identify and quantify anisotropic patterns such as the chessboard or the diagonal stripes that the rectilinear PCF [12] missed. Furthermore, we highlighted how different measures of distance, taxicab and uniform, can lead to different quantifications of spatial correlation.

We extended the calculation of appropriate PCFs to deal with exclusion processes on the other regular spatial tessellations in two dimensions as well as the cubic lattice in three dimensions. We derived the triangle PCF, hexagon PCF, cube taxicab PCF, and cube uniform PCF. These are the first PCFs defined specifically for these discrete lattice types. Finally, we derived a comprehensive PCF for any kind of discrete domain, BC, and metric, which we referred to as the general PCF. The method can be computationally expensive, however, it allows complete freedom in defining a suitable PCF for more complex cases, including those for which recognizing spatial correlation by eye becomes less intuitive.

All of our PCFs are designed for a single species of agents. However, in many applications the agents are divided into multiple species and it can be important to distinguish between different types of spatial correlations: either within agents of the same species (autocorrelation) or comparing the position of agents of different species (cross-correlation). Dini *et al.* [42] have recently investigated correlation in multiple species by using the Rectilinear PCF. We believe a similar approach to that of Dini *et al.* [42] can be applied to our isotropic PCFs to quantify heterogenous correlations, yet it lies beyond the scope of this paper and, as such, we will tackle it in a future publication. The isotropic PCFs that we defined in this paper will be important in further studies and applications. In particular, our functions can improve previous studies (see Ref. [43], for example) which have used PCF as an efficient summary statistics to infer model parameters.

ACKNOWLEDGMENT

The authors would like to thank the CMB-CNCB preprint club for constructive and helpful comments on a preprint of this paper. J.P.O. acknowledges support from the SWBio DTP.

[1] M. S. Steinberg, Adhesion in development: An historical overview, Dev. Biol. **180**, 377 (1996).

[2] A. Cavagna, A. Cimarelli, I. Giardina, G. Parisi, R. Santagati, F. Stefanini, and M. Viale, Scale-free correlations in starling flocks, Proc. Natl. Acad. Sci. U.S.A. **107**, 11865 (2010).

[3] J. E. F. Green, S. L. Waters, J. P. Whiteley, L. Edelstein-Keshet, K. M. Shakesheff, and H. M. Byrne, Nonlocal models for the formation of hepatocyte–stellate cell aggregates, J. Theor. Biol. **267**, 106 (2010).

[4] R. J. Thomas, A. Bennett, B. Thomson, and K. M. Shakesheff, Hepatic stellate cells on poly (dl-lactic acid) surfaces control the formation of 3D hepatocyte coculture aggregates *in vitro*, eCells Mater. **11**, 16 (2005).

[5] F. Chalub, Y. Dolak-Struss, P. Markowich, D. Oelz, C. Schmeiser, and A. Soreff, Model hierarchies for cell aggregation by chemotaxis, Math. Models Methods Appl. Sci. **16**, 1173 (2006).

[6] H. G. Othmer and A. Stevens, Aggregation, blowup, and collapse: The ABC's of taxis in reinforced random walks, SIAM J. Appl. Math. **57**, 1044 (1997).

[7] A. Stevens, A stochastic cellular automaton modeling gliding and aggregation of myxobacteria, SIAM J. Appl. Math. **61**, 172 (2000).

[8] H. Murakawa and H. Togashi, Continuous models for cell-cell adhesion, J. Theor. Biol. **374**, 1 (2015).

[9] M. J. Simpson, B. J. Binder, P. Haridas, B. K. Wood, K. K. Treloar, D. McElwain, and R. E. Baker, Experimental and modeling investigation of monolayer development with clustering, Bull. Math. Biol. **75**, 871 (2013).

[10] K. K. Treloar, M. J. Simpson, P. Haridas, K. J. Manton, D. I. Leavesley, D. S. McElwain, and R. E. Baker, Multiple types of data are required to identify the mechanisms influencing the spatial expansion of melanoma cell colonies, BMC Syst. Biol. **7**, 137 (2013).

[11] E. Hinde, F. Cardarelli, M. A. Digman, and E. Gratton, *In vivo* pair correlation analysis of EGFP intranuclear diffusion reveals DNA-dependent molecular flow, Proc. Natl. Acad. Sci. U.S.A. **107**, 16560 (2010).

[12] B. J. Binder and M. J. Simpson, Quantifying spatial structure in experimental observations and agent-based simulations using pair-correlation functions, Phys. Rev. E **88**, 022705 (2013).

[13] A. Deutsch and S. Dormann, *Cellular Automaton Modeling of Biological Pattern Formation: Characterization, Applications, and Analysis* (Springer Science & Business Media, Berlin, 2007).

[14] R. J. Friedman, D. S. Rigel, and A. W. Kopf, Early detection of malignant melanoma: The role of physician examination and self-examination of the skin, CA. Cancer J. Clin. **35**, 130 (1985).

[15] M. A. Weinstock, Early detection of melanoma, JAMA **284**, 886 (2000).

[16] A. F. G. Bourke and N. R. Franks, *Social Evolution in Ants* (Princeton University Press, Princeton, NJ, 1995).

[17] M. J. Keeling, The effects of local spatial structure on epidemiological invasions, Proc. R. Soc., Ser. B, Biol. Sc., Lond. **266**, 859 (1999).

[18] P. J. Diggle, J. Besag, and T. J. Gleaves, Statistical analysis of spatial point patterns by means of distance methods, Biometrics **32**, 659 (1976).

[19] J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan, *Statistical Analysis and Modeling of Spatial Point Patterns* (John Wiley & Sons, New York, 2008).

[20] D. J. G. Agnew, J. E. F. Green, T. M. Brown, M. J. Simpson, and B. J. Binder, Distinguishing between mechanisms of cell aggregation using pair-correlation functions, J. Theor. Biol. **352**, 16 (2014).

[21] B. J. Binder and M. J. Simpson, Spectral analysis of pair-correlation bandwidth: Application to cell biology images, R. Soc. Open Sci. **2**, 140494 (2015).

[22] B. J. Binder, J. F. Sundstrom, J. M. Gardner, V. Jiranek, and S. G. Oliver, Quantifying two-dimensional filamentous and invasive growth spatial patterns in yeast colonies, Plos Comput. Biol. **11**, e1004070 (2015).

[23] E. J. Hackett-Jones, K. J. Davies, B. J. Binder, and K. A. Landman, Generalized index for spatial data sets as a measure of complete spatial randomness, Phys. Rev. E **85**, 061908 (2012).

[24] N. A. Bahcall and R. M. Soneira, The spatial correlation function of rich clusters of galaxies, Astrophys. J. **270**, 20 (1983).

[25] A. Donev, S. Torquato, and F. H. Stillinger, Pair correlation function characteristics of nearly jammed disordered and ordered hard-sphere packings, Phys. Rev. E **71**, 011105 (2005).

[26] W. R. Young, A. J. Roberts, and G. Stuhne, Reproductive pair correlations and the clustering of organisms, Nature **412**, 328 (2001).

[27] R. N. Binny, A. James, and M. J. Plank, Collective cell behavior with neighbor-dependent proliferation, death, and directional bias, Bull. Math. Biol. **78**, 2277 (2016).

[28] M. Raghib, N. A. Hill, and U. Dieckmann, A multiscale maximum entropy moment closure for locally regulated space–time point process models of population dynamics, J. Math. Biol. **62**, 605 (2011).

[29] D. J. Bone, H. A. Bachor, and R. J. Sandeman, Fringe-pattern analysis using a 2D Fourier transform, Appl. Optics. **25**, 1653 (1986).

[30] M. Takeda, H. Ina, and S. Kobayashi, Fourier-transform method of fringe-pattern analysis for computer-based topography and interferometry, J. Opt. Soc. Am. **72**, 156 (1982).

[31] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.97.062104 for the normalizations of the pair correlation functions that were omitted in the main text, as well as a summary of the normalization formulas. All the MATLAB codes which accompany the paper can be found as online supplementary material.

[32] Q. I. Ouyang and H. L. Swinney, Transition from a uniform state to hexagonal and striped turing patterns, Nature **352**, 610 (1991).

[33] S. Y. Bhide and S. Yashonath, Dependence of the self-diffusion coefficient on the sorbate concentration: A two-dimensional lattice gas model with and without confinement, J. Chem. Phys. **111**, 1658 (1999).

[34] J. D. Murray, *Mathematical Biology: I. An Introduction*, Vol. 17 (Springer Science & Business Media, Berlin, 2007).

[35] P. Ball, Forging patterns and making waves from biology to geology: A commentary on Turing (1952); The chemical basis of morphogenesis, Phil. Trans. R. Soc. B **370**, 20140218 (2015).

[36] S. Kondo, An updated kernel-based turing model for studying the mechanisms of biological pattern formation, J. Theor. Biol. **414**, 120 (2017).

[37] M. J. Simpson, K. A. Landman, and B. D. Hughes, Multi-species simple exclusion processes, Phys. A **388**, 399 (2009).

[38] C. A. Yates, R. E. Baker, R. Erban, and P. K. Maini, Going from microscopic to macroscopic on nonuniform growing domains, Phys. Rev. E **86**, 021921 (2012).

[39] R. J. H. Ross, C. A. Yates, and R. E. Baker, Inference of cell–cell interactions from population density characteristics and cell trajectories on static and growing domains, Math. Biosci. **264**, 108 (2015).

[40] R. E. Baker, C. A. Yates, and R. Erban, From microscopic to macroscopic descriptions of cell migration on growing domains, Bull. Math. Biol. **72**, 719 (2010).

[41] A. P. Browning, S. McCue, and M. J. Simpson, A Bayesian computational approach to explore the optimal duration of a cell proliferation assay, Bull. Math. Biol. **79**, 1888 (2017).

[42] S. Dini, B. J. Binder, and J. E. F. Green, Understanding interactions between populations: Individual-based modeling and quantification using pair correlation functions, J. Theor. Biol. **439**, 50 (2018).

[43] S. T. Johnston, M. J. Simpson, D. L. S. McElwain, B. J. Binder, and J. V. Ross, Interpreting scratch assays using pair density dynamics and approximate Bayesian computation, Op. Bio. **4**, 140097 (2014).

Supplementary materials to accompany: "Pair correlation functions for identifying spatial correlation in discrete domains"

Enrico Gavagnin*, Jennifer P. Owen* and Christian A. Yates

Department of Mathematical Sciences,

University of Bath, Claverton Down, Bath, BA2 7AY, UK

This document contains the supplementary materials which accompany the paper of Gavagnin, Owen, and Yates [2018]. In Section S.1 a full derivation of the normalisation factor for the Square Uniform PCF using non-periodic boundary conditions (BC) is provided. In Section S.2 we report the expressions of the normalisations for the Triangle PCF and Hexagon PCF, using non-periodic BC. Section S.3 contains the derivations of the normalisations for the Cube Taxicab PCF and the Cube Uniform PCF under non-periodic BC. This supplementary document concludes with a summary of the normalisation factors for all of our PCFs in Section S.4. We refer the reader to the main text for the definitions of all of our PCFs.

*These authors contributed equally to this work. Corresponding authors: EG: e.gavagnin@bath.ac.uk, JO: j.owen@bath.ac.uk

S.1 Derivation of the normalisation factor of the Square Uniform PCF under non-periodic BC

In this section we derive the normalisation factor for the Square Uniform PCF, defined in Section 3 in the main text, under non-periodic BC. This corresponds to computing $\mathbb{E}[\bar{c}^n_\infty(m)]$ of equation (9) of the main text. The approach that we use is similar to the one in Section 3.2. In particular, by using the equation (19) of the main text and the expression of $s^p_\infty(m)$ given by equation (22b) we only need to determine the expression of the reminder defined as

$$r_\infty(m) = s^p_\infty(m) - s^n_\infty(m). \tag{S.1}$$

We define the set of pairs of sites separated by distance $m \in \mathcal{D}^n_\infty$ that cross the x boundary (horizontal axis) or y boundary (vertical axis), respectively, as

$$P^x_\infty(m) = \big\{(\boldsymbol{a},\boldsymbol{b}) \in S^p_\infty(m) \;\big|\; |y_a - y_b| > L_y - |y_a - y_b|\big\}, \tag{S.2a}$$
$$P^y_\infty(m) = \big\{(\boldsymbol{a},\boldsymbol{b}) \in S^p_\infty(m) \;\big|\; |x_a - x_b| > L_x - |x_a - x_b|\big\}, \tag{S.2b}$$

where $S^p_\infty(m)$ is defined as in equation (6) of the main text. Within these sets, P^x_∞ and P^y_∞, we select those separated by $k \in \{1, ...m\}$ rows or columns, respectively. We define these subsets as:

$$P^x_\infty(m,k) = \big\{(\boldsymbol{a},\boldsymbol{b}) \in P^x_\infty(m) \;\big|\; L_y - |y_a - y_b| = k\big\}, \tag{S.3a}$$
$$P^y_\infty(m,k) = \big\{(\boldsymbol{a},\boldsymbol{b}) \in P^y_\infty(m) \;\big|\; L_x - |x_a - x_b| = k\big\}. \tag{S.3b}$$

Fig. S 1 (a) and (b) show an example of pairs of sites within $P^x_\infty(m,m)$.

By following the same steps as in Section 3.2, we reduce to the expression for the remainder given by

$$\begin{aligned} r_\infty(m) &= |P^x_\infty(m) \cup P^y_\infty(m)| \\ &= \sum_{k=1}^{m} |P^x_\infty(m,k)| + \sum_{k=1}^{m} |P^y_\infty(m,k)| - |P^x_\infty(m) \cap P^y_\infty(m)|\,. \end{aligned} \tag{S.4}$$

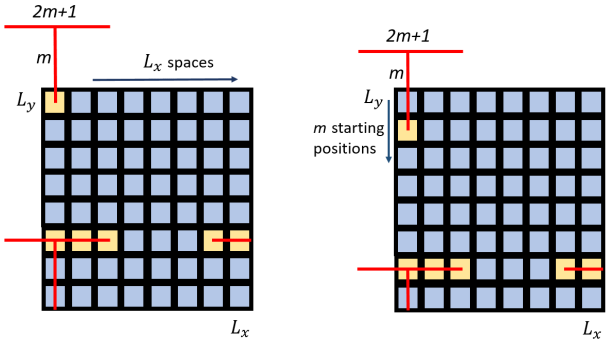
To conclude the computation we derive an expression for the two sums in equation (S.4) and the corresponding equation for the size of the intersection. By counting the contribution of each type of pair (see Fig. S 1 (c) for a visualisation), one can write down the following expressions for the two sums in equation (S.21):

$$\sum_{k=1}^{m} |P^x_\infty(m,k)| = 2(L_x + 2L_x + \dots L_x(m-1)) + m(2m+1)L_x, \tag{S.5a}$$
$$\sum_{k=1}^{m} |P^y_\infty(m,k)| = 2(L_y + 2L_y + \dots L_y(m-1)) + m(2m+1)L_y. \tag{S.5b}$$

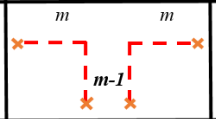
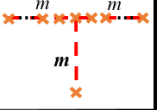
Hence

$$\begin{aligned} \sum_{k=1}^{m} |P^x_\infty(m,k)| + \sum_{k=1}^{m} |P^y_\infty(m,k)| &= 2(L_x + 2L_x +, \dots, L_x(m-1)) + m(2m+1)L_x \\ &\quad + 2(L_y + 2L_y +, \dots, L_y(m-1)) + m(2m+1)L_y \\ &= (L_x + L_y)\left(m(2m+1) + 2\sum_{i=1}^{m-1} i\right) \\ &= (L_x + L_y)\left(m(2m+1) + 2\frac{(m-1)m}{2}\right) \\ &= 3(L_x + L_y)m^2\,. \end{aligned} \tag{S.6}$$

In order to calculate of the residue, $r_\infty(m)$, we need to compute the expression for the size of intersection, $|P^x_\infty(m) \cap P^y_\infty(m)|$, in equation (S.4). This counts the pairs of sites separated by distance m by crossing both the x and y boundaries. Fig. S 2 shows an example of pairs of sites within $P^x_\infty(4) \cap P^y_\infty(4)$.



(a) (b)

k	1	2	...	m-1	m
Illustration of pair types in $P^x_\infty(m,k)$					
$\lvert P^x_\infty(m,k)\rvert$	$2(L_x)$	$2(2L_x)$	...	$2((m-1)L_x)$	$(2m+1)mL_x$

(c)

Fig. S 1: A visualisation of examples of the pairs of sites in $P^x_\infty(m)$. Panels (a) and (b) show different site pairs in $P^x_\infty(m,m)$. Panel (c) shows all the different types of pairs in $P^x_\infty(m,k)$ for $k=1,\ldots,m$ with the corresponding value of $|P^x_\infty(m,k)|$.For example, for each of the L_x columns, each site in the rows $\{L_y-m+1, L_y-m+2,\ldots,L_y\}$ has $2m+1$ corresponding sites at distance m separated by m rows and reached by crossing the horizontal boundary. Therefore $|P^x_\infty(m,m)|=(2m+1)mL_x$.

The number in each site corresponds to the number of sites at distance $m=4$, reached by crossing both the x and y boundaries. So $|P^x_\infty(4)\cap P^y_\infty(4)|$ is the sum of all the numbers in the coloured sites multiplied by two to account for the second corner region.

Extrapolating, for any value of m, the number of pairs of sites that cross the two boundaries is exactly

$$\begin{aligned}
|P^x_\infty(m)\cap P^y_\infty(m)| &= 2\sum_{i=1}^{m}\sum_{j=i}^{m+i-1} j \\
&= \sum_{i=1}^{m}[(m+i)(m+i-1)-i(i-1)] \\
&= \sum_{i=1}^{m}[m^2+2mi-m] \\
&= m^3-m^2+m^2(m+1) \\
&= 2m^3. \qquad \text{(S.7)}
\end{aligned}$$

By substituting equations (S.7) and (S.6) into equation (S.4) we obtain the expression for the remainder $r_\infty(m)$. By rearranging equation (S.1) we determine $s^n_1(m)$, which we then substitute into equation (19) of the main text to obtain the exact expression for the normalisation in the non-periodic case. This is given by

$$\mathbb{E}\left[c^{U,n}_\infty(m)\right] = \left(\frac{N}{L_xL_y}\right)\left(\frac{N-1}{L_xL_y-1}\right)\left(4mL_xL_y-3(L_x+L_y)m^2+2m^3\right). \qquad \text{(S.8)}$$

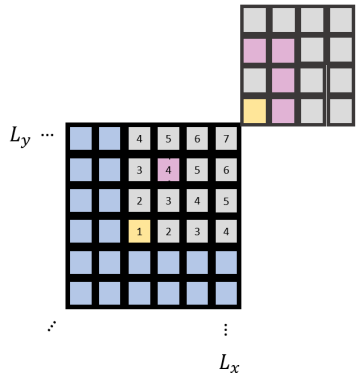


Fig. S 2: Examples of pairs in $P^x_\infty(4) \cap P^y_\infty(4)$ on a zoomed-in corner of a larger domain. Sites in pink and yellow are distance $m = 4$ away from other sites in pink and yellow respectively. The number in each site correspond to the number of sites at distance $m = 4$, reached by crossing the x and y boundaries.

S.2 Normalisation factor for the non-periodic Triangle PCF and Hexagon PCF

Here we provide the expressions of the normalisation factors for the Triangle PCF and the Hexagon PCF defined in Section 5, under non-periodic BC.

In both cases, the normalisation can be computed by using equation (19) of the main text. Hence, we only need to provide the expression of the terms $s^n_{tri}(m)$, for the Triangle PCF, and $s^n_{hex}(m)$, for the Hexagon PCF. These correspond to the exact numbers of site pairs at distance m, assuming non-periodic BC, in the two types of tessellations (see Section 5 of the main text for more details on the definition of the tessellated domain). We omit the derivations of the following expressions which can be obtained with similar steps to the square lattice cases. All the following expressions have been confirmed for a wide range of values of L_x, L_y and m by applying the General PCF (see Section 6 of the main text).

Number of site pairs on a triangular tessellation with non-periodic BC

$$s^n_{tri}(1) = 3\frac{L_x L_y}{2} - \frac{L_x}{2} - L_y\,, \tag{S.9a}$$

$$s^n_{tri}(2) = 3L_x L_y - 2L_x - 4L_y + 2\,, \tag{S.9b}$$

$$\begin{aligned} s^n_{tri}(m) =& 3m\frac{L_x L_y}{2} - \frac{L_x m^2}{2} \qquad\qquad (\text{for } m \geq 3) \\ &+ Ly\left[2k_6(k_6 - 2k_3 + 1) + k_3(m-6) - m^2 + m - 2\right] \\ &+ \frac{1}{3}(m-1)(m^2 - 2m + 6) \\ &- \frac{1}{3}(k_7+1)(20k_7^2 + 37k_7 + 12) \\ &- (m - 7 - 4k_7)(k_7+1)(m + k_7 - 2)\,, \end{aligned} \tag{S.9c}$$

where $k_j = \left\lfloor \frac{m-j}{4} \right\rfloor$.

Number of site pairs on a hexagonal tessellation with non-periodic BC

$$s^n_{hex}(m) = 3mL_xL_y - \frac{1}{4}\left(7m^2 + k\right)L_x - 2m^2L_y + \frac{11m^3}{12} - \frac{(2-3k)m}{12}\,, \tag{S.10}$$

where $k = m \pmod 2$.

S.3 Derivation of the normalisation factor of the Cube Taxicab PCF and Cube Uniform PCF under non-periodic BC

Consider a system of agents on a three-dimensional square lattice of size $L_x \times L_y \times L_z$, with lattice step Δ and with the exclusion property that, at any given time, each lattice site can be occupied by at most one agent. If N agents occupy the domain, then the occupancy of the lattice can be represented by a matrix M:

$$M_{xyz} = \begin{cases} 0 \text{ if } (x,y,z) \text{ is vacant,} \\ 1 \text{ if } (x,y,z) \text{ is occupied.} \end{cases} \tag{S.11}$$

where

$$N = \sum_{x=1}^{L_x} \sum_{y=1}^{L_y} \sum_{z=1}^{L_z} M_{xyz} \leq L_x L_y L_z. \tag{S.12}$$

Let ψ be the set of all agent pairs in the lattice i.e.

$$\psi = \{(\boldsymbol{a}, \boldsymbol{b}) \in \mathbb{L} \times \mathbb{L} \mid \boldsymbol{a} = (x_a, y_a, z_a),\ \boldsymbol{b} = (x_b, y_b, z_b),\ \boldsymbol{a} \neq \boldsymbol{b},\ M_{x_a,y_a,z_a} = M_{x_b,y_b,z_b} = 1\}, \tag{S.13}$$

where $\mathbb{L} = \{1, \ldots, L_x\} \times \{1, \ldots, L_y\} \times \{1, \ldots, L_z\}$ is the set of all sites in the lattice.

In this Section we derive the normalisations for the Cube Taxicab PCF and the Cube Uniform PCF under non-periodic BC. Using the same notation as in Section 2 of the main text, we define the subsets of agent pairs separated by distance m under non-periodic BC as

$$C_1^n(m) = \{(\boldsymbol{a}, \boldsymbol{b}) \in \psi \mid \|\boldsymbol{a} - \boldsymbol{b}\|_1 = m\}, \quad m \in \mathcal{D}_1^n, \tag{S.14a}$$
$$C_\infty^n(m) = \{(\boldsymbol{a}, \boldsymbol{b}) \in \psi \mid \|\boldsymbol{a} - \boldsymbol{b}\|_\infty = m\}, \quad m \in \mathcal{D}_\infty^n, \tag{S.14b}$$

for the taxicab and uniform metric respectively, where $\mathcal{D}_1^n = \mathcal{D}_\infty^n = \{1, 2, \ldots \max\{L_x, L_y, L_z\} - 1\}$. Using the definitions of the uniform and taxicab metrics, we can express these sets as:

$$C_1^n(m) = \{(\boldsymbol{a}, \boldsymbol{b}) \in \psi \mid |x_a - x_b| + |y_a - y_b| + |z_a - z_b| = m\}, \tag{S.15a}$$
$$C_\infty^n(m) = \{(\boldsymbol{a}, \boldsymbol{b}) \in \psi \mid \max\{|x_a - x_b|, |y_a - y_b|, |z_a - z_b|\} = m\}. \tag{S.15b}$$

Similarly we define the subsets of agent pairs separated by distance m under periodic BC as

$$\begin{aligned} C_1^p(m) = \Big\{(\boldsymbol{a}, \boldsymbol{b}) \in \psi \mid & \min\{|x_a - x_b|, L_x - |x_a - x_b|\} \\ & + \min\{|y_a - y_b|, L_y - |y_a - y_b|\} \\ & + \min\{|z_a - z_b|, L_y - |y_a - y_b|\} = m\Big\}, \quad m \in \mathcal{D}_1^p, \end{aligned} \tag{S.16a}$$

$$\begin{aligned} C_\infty^p(m) = \Big\{(\boldsymbol{a}, \boldsymbol{b}) \in \psi \mid \max\Big\{ & \min\{|x_a - x_b|, L_x - |x_a - x_b|\}, \\ & \min\{|y_a - y_b|, L_y - |y_a - y_b|\}, \\ & \min\{|z_a - z_b|, L_z - |z_a - z_b|\}\Big\} = m\Big\}, \quad m \in \mathcal{D}_\infty^p, \end{aligned} \tag{S.16b}$$

where $\mathcal{D}_1^p = \mathcal{D}_\infty^p = \{1, 2, \ldots \min\{\lfloor \frac{L_x}{2} \rfloor, \lfloor \frac{L_y}{2} \rfloor, \lfloor \frac{L_z}{2} \rfloor\}\}$. The schematics in Fig. 12 of the main text represent examples of sites separated by distance $m = 1$ using the taxicab (a) and uniform (b) metrics in three dimensions. The formulae for the number of pairs of sites separated by distance m under periodic BC are given in Section 5.2 of the main text. Here we compute the numbers of pairs of sites separated by distance m under non-periodic BC, *i.e* $s_1^n(m)$ and $s_\infty^n(m)$, which we can then substitute into the equation (19) in the main text to obtain the normalisation factors of the corresponding PCFs.

Similarly to the two-dimensional cases, we focus on computing an expression the remainder given by

$$r_d(m) = s_d^p(m) - s_d^n(m), \tag{S.17}$$

where $d = 1, \infty$. The expression of $s_d^n(m)$ will then follow by rearrangement.

Let us define the set of pairs of sites separated by distance $m \in \mathcal{D}_1^n$ that cross the y-z plane boundary (*i.e.* $x = 0$), x-z plane boundary (*i.e.* $y = 0$) and x-y plane boundary (*i.e.* $z = 0$), respectively as

$$P_1^{yz}(m) = \big\{(\boldsymbol{a}, \boldsymbol{b}) \in S_1^p(m) \;\big|\; |x_a - x_b| > L_x - |x_a - x_b|\big\}, \tag{S.18a}$$
$$P_1^{xz}(m) = \big\{(\boldsymbol{a}, \boldsymbol{b}) \in S_1^p(m) \;\big|\; |y_a - y_b| > L_y - |y_a - y_b|\big\}, \tag{S.18b}$$
$$P_1^{xy}(m) = \big\{(\boldsymbol{a}, \boldsymbol{b}) \in S_1^p(m) \;\big|\; |z_a - z_b| > L_z - |z_a - z_b|\big\}. \tag{S.18c}$$

For $P_1^{yz}(m)$, $P_1^{xz}(m)$ and $P_1^{xy}(m)$ let us consider those pairs of sites separated by $k \in \{1, ...m\}$ sites in the corresponding orthogonal direction, *i.e.* $x-$, $y-$ and $z-$direction, respectively. We define these subsets as:

$$P_1^{yz}(m,k) = \big\{(\boldsymbol{a}, \boldsymbol{b}) \in P_1^{yz}(m) \big| \quad L_x - |x_a - x_b| = k\big\}, \tag{S.19a}$$
$$P_1^{xz}(m,k) = \big\{(\boldsymbol{a}, \boldsymbol{b}) \in P_1^{xz}(m) \big| \quad L_y - |y_a - y_b| = k\big\} \tag{S.19b}$$
$$P_1^{xy}(m,k) = \big\{(\boldsymbol{a}, \boldsymbol{b}) \in P_1^{xy}(m) \big| \quad L_z - |z_a - z_b| = k\big\}. \tag{S.19c}$$

Notice that we have $P_d^{yz}(m) = \bigcup_{k=1}^m P_d^{yz}(m,k)$, $P_d^{xz}(m) = \bigcup_{k=1}^m P_d^{xz}(m,k)$ and $P_d^{xy}(m) = \bigcup_{k=1}^m P_d^{xy}(m,k)$. Figs. S.3 (a)-(b) provide a visualisation of pairs of sites within $P_1^{xy}(m,m)$. In Fig. S.3 (c) we show the possible types of pairs of sites in $P_1^x(m,k)$, for $k = 1, \ldots, m$. By definitions (S.14) and (S.16) we have that

$$S_d^p(m) \setminus S_d^n(m) = P_d^{yz}(m) \cup P_d^{xz}(m) \cup P_d^{xy}(m)\,, \tag{S.20}$$

where $d = 1, \infty$. Hence, we obtain

$$\begin{aligned} r_d(m) =& |P_d^{yz}(m) \cup P_d^{xz}(m) \cup P_d^{xy}(m)| \\ =& \sum_{k=1}^m |P_d^{yz}(m,k)| + \sum_{k=1}^m |P_d^{xz}(m,k)| + \sum_{k=1}^m |P_d^{xy}(m,k)| \\ & - |P_d^{yz}(m) \cap P_d^{xy}(m)| - |P_d^{xz}(m) \cap P_d^{yz}(m)| - |P_d^{xz}(m) \cap P_d^{xy}(m)| \\ & + |P_d^{yz}(m) \cap P_d^{xz}(m) \cap P_d^{xy}(m)|\,, \end{aligned} \tag{S.21}$$

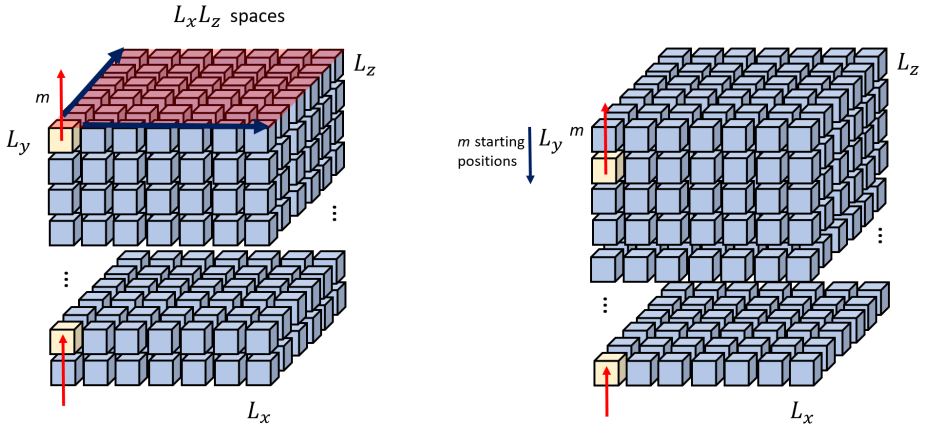
for $d = 1, \infty$.

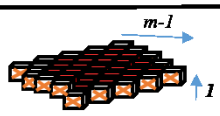
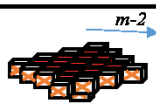
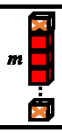
By counting the contribution of each type of pair (see Fig. S.3 (c) for a visualisation), one can write down the following expressions for the first three sums in equation (S.21):

$$\sum_{k=1}^m |P_1^{ij}(m,k)| = L_i L_j \Big(4\,(m - 1 + 2(m-2) + 3(m-3) + ... + m - 1) + m\Big), \tag{S.22a}$$
$$\sum_{k=1}^m |P_\infty^{ij}(m,k)| = L_i L_j \Big(8m\,(m - 1 + (m-2) + (m-3) + ... + 1) + (2m+1)^2\Big), \tag{S.22b}$$

where $ij = yz, xz, xy$. Hence for the case of taxicab metric we have

$$\begin{aligned} \sum_{k=1}^m |P_1^{yz}(m,k)| + \sum_{k=1}^m |P_1^{xz}(m,k)| + \sum_{k=1}^m |P_1^{xy}(m,k)| & \\ =& \Big(4 \sum_{i=1}^{m-1} i(m-i)\Big)(L_x L_y + L_y L_z + L_z L_x) \\ =& \frac{2}{3} m(m-1)(m+1)(L_x L_y + L_y L_z + L_z L_x) \\ =& \frac{1}{3}(2m^3 + m)(L_x L_y + L_y L_z + L_z L_x)\,; \end{aligned} \tag{S.23}$$



k	1	2	...	m-1	m
Illustration of pair types in $P_1^{xz}(m,k)$					
$\lvert P_1^{xz}(m,k)\rvert$	$4(m-1)L_xL_z$	$4(m-2)2L_xL_z$	...	$4(m-1)L_xL_z$	mL_xL_z

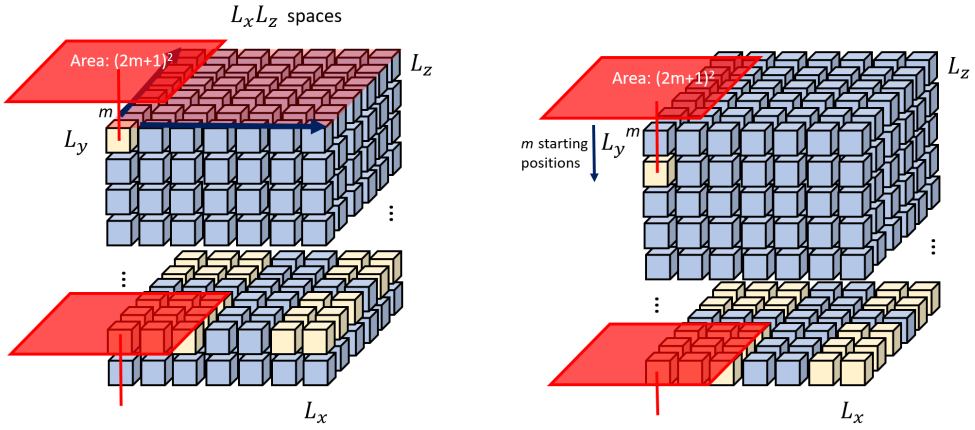
(c)

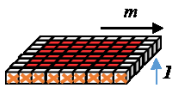
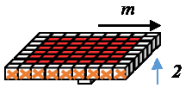
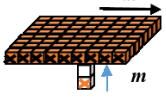
Fig. S 3: A visualisation of the pairs of sites in $P_1^{xz}(m)$. Panels (a) and (b) show two different site pairs in $P_1^{xz}(m,m)$. Panel (c) shows all the different types of pairs in $P_1^{xz}(m,k)$ for $k = 1,\dots,m$ with the corresponding value of $|P_1^{xz}(m,k)|$. For example, each site in the top m horizontal planes, *i.e.* $\{y = L_y - m + 1, y = L_y - m + 2, \dots, y = L_y\}$, has a single corresponding site at distance m separated by m horizontal planes and 0 vertical planes, which is reached by crossing the x-z boundary. Therefore $|P_1^{xz}(m,m)| = mL_xL_y$. The axis orientation is chosen to be consistent with the two-dimensional case.

and similarly for the uniform metric:

$$\begin{aligned}\sum_{k=1}^{m}|P_\infty^{yz}(m,k)| + \sum_{k=1}^{m}|P_\infty^{xz}(m,k)| + \sum_{k=1}^{m}|P_\infty^{xy}(m,k)| \\ =&\left((2m+1)^2 + 8m\sum_{i=1}^{m-1} i\right)(L_xL_y + L_yL_z + L_zL_x) \\ =&m(8m^2+1)(L_xL_y + L_yL_z + L_zL_x)\,. \end{aligned} \tag{S.24}$$

We now focus on deriving an expression for the unknown expressions in (S.21), that is the size of intersections, $P_d^{xy}(m) \cap P_d^{yz}(m)$, $P_d^{xy} \cap P_d^{xz}(m)$ and $P_d^{xy} \cap P_d^{xz}(m)$. These sets consist of pairs of sites separated by distance m that cross *at least* two of the three y-z, x-z, x-y boundaries. Examples of pairs of sites in $P_1^{xz}(m) \cap P_1^{yz}(m)$ are illustrated in Fig. S 5 (a) and Fig. S 5 (b). We show that $P_1^{xz}(m) \cap P_1^{yz}(m)$ consists of all of the pairs of sites located in the top left and right edges (right demonstrated only) which cross the x-z and y-z boundaries of each of the L_z vertical planes making up the cube (see Fig. S 5 (b)). The sites that comprise pairs in $P_1^{xz}(m) \cap P_1^{yz}(m)$ can both be in the same x-y plane or they can be in neighbouring (up to $m-2$) planes away. The number of pairs in these sets can then be calculated by addition, enumerating possible pairs as in Fig. 6 in the main text for taxicab case and Fig. S 2 for the



k	1	2	...	m
Illustration of pair types in $P_\infty^{xz}(m,k)$				
$\lvert P_\infty^{xz}(m,k)\rvert$	$8m(m-1)L_xL_z$	$8m(m-2)L_xL_z$	...	$(2m+1)^2 L_xL_z$

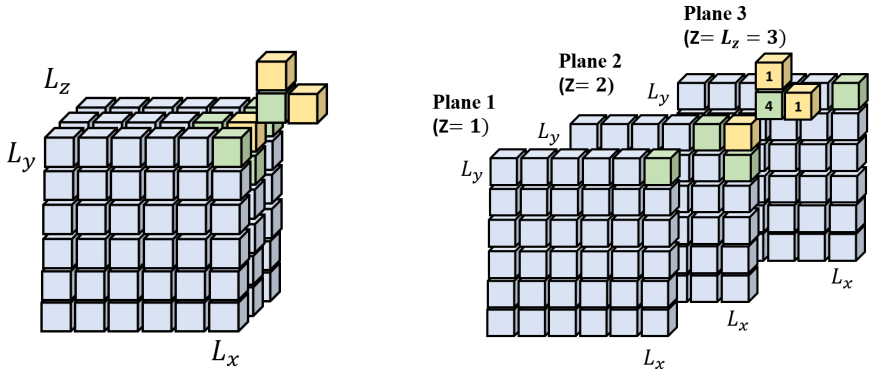
(c)

Fig. S 4: A visualisation of the pairs of sites in $P_\infty^{xz}(m)$. Panels (a) and (b) show two different site pairs in $P_\infty^{xz}(m,m)$. Panel (c) shows all the different types of pairs in $P_\infty^{xz}(m,k)$ for $k=1,\ldots,m$ with the corresponding value of $|P_\infty^{xz}(m,k)|$. For example, each site in the top m horizontal planes, *i.e.* $\{y=L_y-m+1, y=L_y-m+2,\ldots,y=L_y\}$, has $(2m+1)^2$ sites at distance m separated by m horizontal planes which are reached by crossing the x-z boundary. Therefore $|P_\infty^{xz}(m,m)|=(2m+1)^2L_xL_z$.

uniform case respectively to give the following. For any $s,t,u\in\{x,y,z\}$, such that $(s,t)\neq(t,u)$

$$\begin{aligned}|P_1^{st}(m)\cap P_1^{tu}(m)| =& 2L_t\sum_{i=1}^{m} i^2(m-i)\\ =&\frac{1}{6}(m-1)m^2(m+1)L_t, \qquad \text{(S.25)}\\ |P_\infty^{st}(m)\cap P_\infty^{tu}(m)| =& 2L_t\left((2m+1)m^3+2\sum_{i=1}^{m-1}\sum_{j=1}^{m-1} ij\right)\\ =&m^2(5m^2+1)L_t. \qquad \text{(S.26)}\end{aligned}$$

Finally, we compute $|P_d^{yz}(m)\cap P_d^{xz}(m)\cap P_d^{xy}(m)|$, the number of pairs of sites of distance m that cross the y-z, x-z and x-y plane boundaries. Notice that if $m<3$ no pairs can be connected by crossing all three boundaries, which means that $P_1^{yz}(m)\cap P_1^{xz}(m)\cap P_1^{xy}(m)$ is the empty set. Fig. S 6 is an illustration of pairs of sites in $P_1^{yz}(4)\cap P_1^{xz}(4)\cap P_1^{xy}(4)$. These consist of the pairs of sites in four of the eight corners of the cube. In the case of the taxicab metric, the number of pairs of sites in $P_1^{yz}(m)\cap P_1^{xz}(m)\cap P_1^{xy}(m)$ can be calculated by first considering the site at the very corner of the cube (marked in green in Fig. S 6). There is exactly $Tr(1)=1$ site at the very corner of the cube where $Tr(x)$ refers to the xth triangular number. This site is distance m from $Tr(m-2)$ sites (yellow sites outside of the domain) by crossing the y-z, x-z and x-y plane boundaries. Next we consider the nearest neighbouring sites to the very corner of the cube. There are $Tr(2)=3$ of these sites (marked in yellow). Each of the these sites are distance m from $Tr(m-3)$ sites by crossing the y-z, x-z and x-y plane boundaries (marked in yellow outside of the



(a) (b)

Fig. S 5: Examples of pairs of sites separated by distance $m = 3$ that cross both the x-z and y-z plane boundaries, *i.e.* pairs of sites in $P_1^{xz}(3) \cap P_1^{yz}(3)$. Green and yellow sites inside of the cube domain are distance $m = 3$ from green and yellow sites outside of the domain respectively. Panel (a) illustrates an example three-dimensional cube that can be split into L_z vertical planes, panel (b) illustrates how the x-y plane $z = 2$ contributes sites that are distance m from other sites by crossing both the x-z and y-z boundaries.

domain). We continue this method to count site pairs until we consider sites that separated by $m - 3$ sites from the corner site (green site in Fig. S 6). Hence we deduce that:

$$\begin{aligned} |P_1^{yz}(m) \cap P_1^{xz}(m) \cap P_1^{xy}(m)| =& 4 \sum_{i=1}^{m-2} Tr(i) Tr(m-i-1) \\ =& 4 \sum_{i=1}^{m-2} \frac{i(i+1)}{2} \cdot \frac{(m-i-1)(m-i)}{2}, \\ =& \frac{m^5}{30} - \frac{m^3}{6} + \frac{2m}{15}. \end{aligned} \tag{S.27}$$

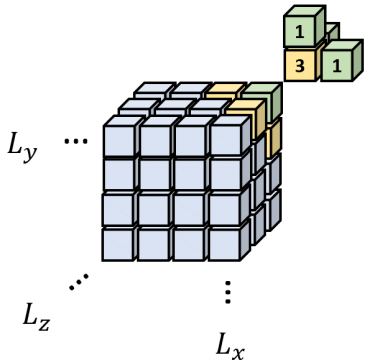


Fig. S 6: Examples of pairs of sites separated by distance $m = 4$ that cross the x-y, y-z and x-z boundaries, *i.e.* pairs of sites in $P_1^{xz}(4) \cap P_1^{yz}(4) \cap P_1^{xy}(4)$. Green and yellows sites inside of the cube domain are distance $m = 3$ from green and yellow sites outside of the domain. The numbers correspond to the number of distance $m = 4$ neighbours each site has which lie in $P_1^{xz}(4) \cap P_1^{yz}(4) \cap P_1^{xy}(4)$.

With a similar argument, one can obtain the corresponding expression by using the uniform metric

which reads

$$|P_{\infty}^{yz}(m) \cap P_{\infty}^{xz}(m) \cap P_{\infty}^{xy}(m)| = m^3(3m^2+1)\,. \tag{S.28}$$

The formulae (S.27) and (S.28) complete the formulae needed for the computation of the remainder of equation (S.21). This can be used together with Eqs. (38) in the main text to rearrange equation (S.17). The resulting expressions for the numbers of pairs at distance m under non-periodic BC in three dimensions, for the taxicab and uniform metrics, respectively, read

$$\begin{aligned} s^n_{cube_1}(m) =& (2m^2+1)L_xL_yL_z - \frac{1}{3}(2m^3+m)(L_xL_y+L_yL_z+L_zL_x) \\ &+ \frac{m^2(m^2-1)}{6}(L_x+L_y+L_z) - \frac{m^5}{30} + \frac{m^3}{6} - \frac{2m}{15}\,, \end{aligned} \tag{S.29a}$$

and

$$\begin{aligned} s^n_{cube_\infty}(m) =& (12m^2+1)L_xL_yL_z - m(8m^2+1)(L_xL_y+L_yL_z+L_zL_x) \\ &+ m^2(5m^2+1)(L_x+L_y+L_z) - m^3(3m^2+1)\,. \end{aligned} \tag{S.29b}$$

The normalisations for the Cube Taxicab PCF and Cube Uniform PCF under non-periodic BC are obtained by substituting expressions (S.29) into equation (19) in the main text.

S.4 Summary of normalisation factors

In this section we summarise all of the normalisation factors which have been computed in the paper. For all cases considered, the normalisation, $\mathbb{E}[\bar{c}_d(m)]$, is obtained by equation (19) in the main text. In Table 1 we report only the counts of pairs of sites separated by distance m, $s_d(m)$, depending on the metric and the BC used.

PCF (d)	Periodic (s_d^p)	Non-periodic (s_d^n)
Square Taxicab (1)	$2mL_xL_y$	$2mL_xL_y - (L_x+L_y)m^2 + \frac{m^3-m}{3}$
Square Uniform (∞)	$4mL_xL_y$	$4mL_xL_y - 3(L_x+L_y)m^2 + 2m^3$
Triangle (tri)	$\frac{3}{2}mL_xL_y$	for $m=1$: $\quad 3\frac{L_xL_y}{2} - \frac{L_x}{2} - L_y \quad ,$ for $m \geq 3$: $\quad 3m\frac{L_xL_y}{2} - \frac{m^2}{2}L_x + \left[2k_6(k_6-2k_3+1) + k_3(m-6) - m^2 + m - 2\right]L_y$ $+\frac{1}{3}(m-1)(m^2-2m+6) - \frac{1}{3}(k_7+1)(20k_7^2+37k_7+12)$ $-(m-7-4k_7)(k_7+1)(m+k_7-2)\,, \qquad \text{where } k_j = \left\lfloor \frac{m-j}{4} \right\rfloor$
Hexagon (hex)	$3mL_xL_y$	$3mL_xL_y - \frac{1}{4}\left(7m^2+k\right)L_x - 2m^2L_y + \frac{11m^3}{12} - \frac{(2-3k)m}{12}$ where $k = m \pmod 2$
Cube Taxicab ($cube_1$)	$(2m^2+1)L_xL_yL_z$	$(2m^2+1)L_xL_yL_z - \frac{1}{3}(2m^3+m)(L_xL_y+L_yL_z+L_zL_x)$ $+\frac{m^2(m^2-1)}{6}(L_x+L_y+L_z) - \frac{m^5}{30} + \frac{m^3}{6} - \frac{2m}{15}$
Cube Uniform ($cube_\infty$)	$(12m^2+1)L_xL_yL_z$	$(12m^2+1)L_xL_yL_z - m(8m^2+1)(L_xL_y+L_yL_z+L_zL_x)$ $+m^2(5m^2+1)(L_x+L_y+L_z) - m^3(3m^2+1)$

Table 1: Summary of the number of sites pairs, $s_d(m)$, according to the metric and BC used.

References

E. Gavagnin, J.P. Owen, and C.A. Yates. Pair correlation functions for identifying spatial correlation in discrete domains. *Phys. Rev. E*, 97:062104, 2018.

5.3. Conclusions

In this chapter, I presented our work identifying and quantifying spatial correlation in discrete domains. The aim of this paper is to improve the analysis of spatial correlation in discrete domains thus allowing the study of spatial phenomena for on-lattice models. The PCFs designed in this chapter should be considered a set of promising tools for any scenario in which spatial correlation is important. For example, Cutright *et al.* recently cited our paper in the context of microgel packing arrangement [106]. The PCF could also be used to validate and analyse future discrete models of animal pigment patterns. In the context of this thesis, the results of this paper are inspired by and applicable to the next chapter, Chapter 6 wherein I validate my mathematical model of zebrafish stripe formation using these metrics. In particular, I use the PCF to determine the mean width of the stripes generated using my mathematical model of zebrafish stripe formation and compare those with real zebrafish stripes. Furthermore, by altering the parameters accordingly, I compare between the sizes of melanophore spots produced when replicating different spotted mutants such as *shd* and *sbr* (seen in Chapter 3), thus using the PCF to validate my modelling approach.

Chapter 6

A quantitative modelling approach to zebrafish pigment pattern formation

6.1. Overview

In Chapters 3 and 4 we compiled the biological evidence relating to zebrafish stripe formation as well as the benefits and pitfalls of previous mathematical models describing them. In this chapter, we use the biological evidence from Chapter 3 in order to build our own comprehensive mathematical model of zebrafish stripe formation. The chapter comprises the paper written by myself (Jennifer Owen), Christian Yates and Robert Kelsh as published in eLife [28] in 2020 and forms a major part of this thesis. The aim of the paper is to use our model to both determine whether the evidenced cell-cell interaction rules are sufficient to determine pattern formation and to further make and test predictions about the pattern forming process.

In this study we define a comprehensive stochastic on-lattice mathematical model of zebrafish stripe formation whereby we model cells as individual agents that interact, die and differentiate, in a manner determined by the literature. Where the biology is lacking, as in the case of iridophore shape-changes from loose to dense, we deduce a minimal set of rules for these interactions based on the pattern formation of missing-cell mutants. We show that, once implemented, not only can our model generate WT and missing-cell mutant patterns, but also by changing the parameters accordingly, the model can reproduce other mutant patterns too, for example *rose*, *seurat*, *schachbrett* and *choker*. We show that our model generates patterns that qualitatively and quantitatively match real patterns. Thus our results demonstrate the general sufficiency of the biological rules to generate the patterns. Next we use the model in order to make predictions about the system. By changing the position of the initial iridophore interstripe, we are able to demonstrate that the position alone does not determine the orientation of the stripes and interstripes. Furthermore, by analysing mutants *leo*, *leo;shd*, *leo;nac* and *leo;pfe* we are able to derive hypotheses about the functionality of the *leo* gene. Our study demonstrates the capability of mathematical modelling to answer questions about a complex system and will guide future experimental research into zebrafish stripe formation.

6.2. Outline of the chapter

We provide a brief overview of the zebrafish stripe formation process as well as previous mathematical models of the zebrafish pattern formation process in the introduction of this paper. In the materials and methods section we describe our mathematical model of zebrafish stripe formation. In the Results section we demonstrate the capability of the model to generate a mixture of wild-type and mutant patterns, both qualitatively and quantitatively, thus demonstrating sufficiency of the biological rules. Having validated the model, we further use it to explore important mechanisms for controlling pattern formation as well as to hypothesise the functionality of the *leo* gene. We summarise our findings, as well as future work in the Discussion. The supplementary material provides more details about the implementation of our mathematical model as well as

some extra predictive studies using the model.

Appendix B: Statement of Authorship

This declaration concerns the article entitled:			
A quantitative modelling approach to zebrafish pigment pattern formation			
Publication status (tick one) **Draft manuscript** [ ] **Submitted** [ ] **In review** [ ] **Accepted** [ ] **Published** [X]			
Publication details (reference)	Journal: eLife, 9, 062020 Authors: Jennifer P. Owen, Christian A. Yates and Robert N. Kelsh		
Copyright status (tick the appropriate statement) I hold the copyright for this material [ ] Copyright is retained by the publisher, but I have been given permission to replicate the material here [X]			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	All the simulations and the numerical computations have been performed by the author of the thesis (100%). All authors contributed equally to the presentation of the content (50%).		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed	Jennifer Owen	**Date**	07/06/2021

RESEARCH ARTICLE

A quantitative modelling approach to zebrafish pigment pattern formation

Jennifer P Owen*, Robert N Kelsh†, Christian A Yates†*

Department of Biology and Biochemistry and Department of Mathematical Sciences, University of Bath, Claverton Down, Bath, United Kingdom

Abstract Pattern formation is a key aspect of development. Adult zebrafish exhibit a striking striped pattern generated through the self-organisation of three different chromatophores. Numerous investigations have revealed a multitude of individual cell-cell interactions important for this self-organisation, but it has remained unclear whether these known biological rules were sufficient to explain pattern formation. To test this, we present an individual-based mathematical model incorporating all the important cell-types and known interactions. The model qualitatively and quantitatively reproduces wild type and mutant pigment pattern development. We use it to resolve a number of outstanding biological uncertainties, including the roles of domain growth and the initial iridophore stripe, and to generate hypotheses about the functions of *leopard*. We conclude that our rule-set is sufficient to recapitulate wild-type and mutant patterns. Our work now leads the way for further in silico exploration of the developmental and evolutionary implications of this pigment patterning system.

***For correspondence:**
jpo22@bath.ac.uk (JPO);
c.yates@bath.ac.uk (CAY)

†These authors contributed equally to this work

Competing interests: The authors declare that no competing interests exist.

Funding: See page 37

Received: 23 October 2019
Accepted: 21 June 2020
Published: 27 July 2020

Reviewing editor: Sandeep Krishna, National Centre for Biological Sciences-Tata Institute of Fundamental Research, India

Introduction

Pattern formation - the process generating regular features from homogeneity - is a fascinating phenomenon that is as ubiquitous as it is diverse. It is a major aspect of developmental biology, with key exemplars including segmentation within the syncitial blastoderm of fruit flies (***Clark and Peel, 2018***), digit formation in the vertebrate limb (***Tickle, 2006***), and branching patterns in kidney and lung development (***Davies, 2002***).

Another key example, pigment pattern formation, the process generating functional and often beautiful distributions of pigment cells, represents a classic problem in both developmental and mathematical biology. Pigment patterns allow animals to distinguish between individuals within a group and identify those of different species and are an important characteristic for the survival of most animals in wild populations. Pigment patterns are striking. They form rapidly and, in many cases, autonomously, that is, the process relies on self-organisation and not internal body structures. Additionally, they often vary dramatically between even closely related species, therefore recognising similarities and differences in the development of these related species can allow us insight into the evolutionary change. Finally, pigment pattern formation is made experimentally tractable by the self-labelling nature of pigment cells.

The horizontal blue and gold stripes of zebrafish are now one of the best-studied examples of pigment pattern formation, especially at the level of underlying cellular mechanisms (***Singh and Nüsslein-Volhard, 2015***; ***Watanabe and Kondo, 2015***; ***Patterson and Parichy, 2019***). Zebrafish are amenable to observational studies, since all development takes place outside the mother and the skin is transparent. This, combined with the availability of multiple key mutants (affecting, for example, cell-type differentiation and patterning), and the development of innovative in vivo cell ablation and in vitro cell culture techniques, have provided a unique opportunity to investigate the cellular and molecular basis for pigment pattern formation experimentally (***Eom et al., 2015***; ***Budi et al., 2011***; ***Ceinos et al., 2015***; ***Yamanaka and Kondo, 2014***; ***Eom and Parichy, 2017***; ***Hamada et al.,***

2014; *Fadeev et al., 2015*; *Inoue et al., 2014*; *Irion et al., 2014*; *Watanabe et al., 2006*; *Frohnhöfer et al., 2013*; *Parichy et al., 2009*; *Svetic et al., 2007*; *Mellgren and Johnson, 2006*; *Parichy et al., 2000b*; *Iwashita et al., 2006*; *Hirata et al., 2005*; *Kelsh et al., 1996*; *Lister et al., 1999*; *Parichy et al., 2000a*; *Maderspacher and Nüsslein-Volhard, 2003*; *Walderich et al., 2016*; *Patterson et al., 2014*; *McMenamin et al., 2014*; *Patterson and Parichy, 2013*; *Krauss et al., 2014*; *Parichy and Turner, 2003*; *Eom et al., 2012*; *Mahalwar et al., 2016*; *Mahalwar et al., 2014*; *Asai et al., 1999*; *Takahashi and Kondo, 2008*).

The cellular composition of the stripes and how these become assembled has been well-described. Zebrafish generate, over a period of a few weeks and beginning around 3 weeks of age (*Frohnhöfer et al., 2013*; *Parichy et al., 2009*), a robust adult stripe pattern of alternating dark blue stripes and golden interstripes comprised of three different pigment-producing cell types: melanocytes, containing black melanin; xanthophores containing yellow and orange carotenoids and pteridines; and iridescent iridophores, containing guanine crystals within reflective platelets (*Hirata et al., 2003*; *Figure 1A*).

Of the iridophores, there are two types distinguished by their platelet distribution (*Hirata et al., 2003*); type L-iridophores, and type S-iridophores. Only type S-iridophores play a role in stripe formation (type L-iridophores appear later and are likely involved in pattern maintenance [*Hirata et al., 2003*]), and they appear in two different forms. In the light interstripes, S-iridophores appear in a 'dense' arrangement (dense S-iridophores), forming a continuous sheet, whilst in the dark stripes the cells are in a 'loose' arrangement and appear more widely spaced (loose S-iridophores) (*Hirata et al., 2003*; *Fadeev et al., 2015*). Pigment cells are found in the hypodermis below the dermis, and organised as layers of cells consistently stacked in the same order (*Hirata et al., 2003*). Starting from the deepest layer of the hypodermis just above the muscle and moving to the dermis, adult dark stripes consist of consecutive layers of L-iridophores, melanocytes, loose S-iridophores

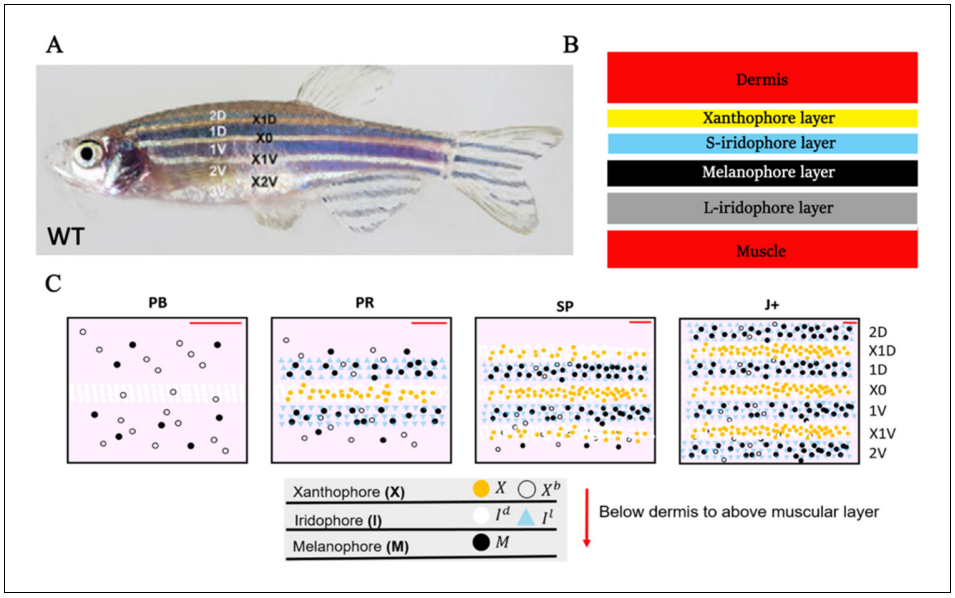


Figure 1. WT stripe composition and development. (**A**) An adult wild type (WT) fish. Stripes and interstripes are labelled according to their order of temporal appearance. X0 is the first interstripe to appear. 1D and 1V (D - Dorsal, V- Ventral) are the first two stripes to appear. X1D and X1V are the next two interstripes to appear and so on. Image reproduced from *Frohnhöfer et al., 2013* and licensed under CC-BY 4.0 (https://creativecommons.org/licenses/by/4.0). (**B**) Summary of pigment cell distribution in adult zebrafish. The cells in the xanthophore, S-iridophore, melanophore and L-iridophore layers consist of xanthophores and xanthoblasts, melanophores, S-iridophores and L-iridophores, respectively. Adapted from *Hirata et al., 2003*. (**C**) Schematic of WT patterns on the body of zebrafish. Stages PB, PR, SP, J+ correspond to developmental stages described in 3. Patterns form sequentially outward from the central interstripe, labelled X0, with additional dorsal stripes and interstripes labelled 1D, X1D, 2D, X2D from the centre (horizontal myoseptum) dorsally outward (similarly, ventral stripes and interstripes are labelled 1V, X1V, 2V, etc).

and xanthophores. Similarly, adult light interstripes are made up of layers of dense S-iridophores and xanthophores (***Figure 1B***; ***Hirata et al., 2003***). The final striped pattern is generated by the self-organisation of xanthophores, melanocytes, loose and dense S-iridophores into the appropriate positions within the hypodermis.

Prior to the initiation of adult stripe formation zebrafish exhibit a larval pigment pattern, formed in the first 5 days of development. Embryonic pigment cells form a distinctive early larval pattern that is essentially complete by 5 days post-fertilisation (dpf) and remains unchanged until metamorphosis. This pattern consists of melanocytes in four stripes (dorsal to the central nervous system, within the horizontal myoseptum, dorsal to the gut and ventrally under the yolk; S-iridophores are found associated with three of these melanocyte stripes [***Parichy et al., 2009***; ***Frohnhöfer et al., 2013***]). Xanthophores lie in a monolayer under the skin, filling the areas between the melanocytes above the CNS and extending ventrally to the level of the gut. Formation of the adult pattern involves replacement of melanocytes and S-iridophores with new cells derived from adult pigment stem cells. Early larval xanthophores dedifferentiate, forming unpigmented xanthoblasts that regain their proliferative ability and proceed to generate the adult xanthophores (***McMenamin et al., 2014***; ***Mahalwar et al., 2014***; ***Parichy et al., 2000b***); an unknown proportion of the latter may derive from de novo production from adult pigment stem cells (***Kelsh et al., 2017***). Xanthophore de-differentiation is complete by 21dpf, and early metamorphic melanocytes appear in a widely scattered distribution between 14 and 21dpf (***Parichy et al., 2009***), thus forming the initial metamorphic pattern (***Figure 1C***, stage PB). A key event in the initiation of adult pattern metamorphosis is the appearance of newly differentiated dense S-iridophores alongside the horizontal myoseptum. In response to the appearance of these S-iridophores, the first adult xanthophores are generated (***Figure 1C***, stage PR) by differentiation from xanthoblasts in this region, thus initiating the first interstripe, X0. Furthermore, metamorphic melanocytes begin to accumulate either side of this central interstripe, marking the first two stripes denoted 1D and 1V (***Figure 1C***, stage PR). Subsequently, S-iridophores proliferate rapidly and spread bidirectionally; at the edges of the interstripes they switch to a more scattered (less tightly-packed) form as they continue to spread dorsally and ventrally. Spreading loose S-iridophores transition back into dense S-iridophores at the locations of the future interstripes X1V and X1D (***Figure 1C***, stages PR to SP) (***Mahalwar et al., 2014***). Once S-iridophores aggregate at the next interstripe, the process starts again, that is, xanthophores differentiate in response to the dense S-iridophores and melanocytes accumulate either side of the new interstripe generating the subsequent stripe. This process of S-iridophore aggregation predetermining future interstripe locations and subsequent delamination in future stripe regions repeats until S-iridophores cover the domain and all stripes (between 4 and 5) and interstripes are fully formed (***Figure 1C***, stage J+).

In addition to the description of pattern development (***Frohnhöfer et al., 2013***), many studies have identified individual patterning mechanisms that contribute to stripe formation, although it is unclear whether these are sufficient to explain pattern formation. Stripe generation is complex and requires many interactions. During pattern metamorphosis, these interactions may determine cell birth (***Mahalwar et al., 2014***), cell death (***Takahashi and Kondo, 2008***), cell migration (***Yamanaka and Kondo, 2014***; ***Takahashi and Kondo, 2008***; ***Patterson et al., 2014***), long-distance communication, through stabilisation of elongated cellular projections (***Eom and Parichy, 2017***; ***Eom et al., 2015***), as well as the shape transitions of S-iridophores (***Fadeev et al., 2015***). During this period, there is also simultaneous two-dimensional domain growth (***Parichy et al., 2009***). The pattern is formed by cell–cell interactions of all three pigment producing cell types: melanocytes, xanthophores and S-iridophores. Without any one of these cell types, pattern formation is disrupted (***Frohnhöfer et al., 2013***; ***Patterson and Parichy, 2013***).

Mathematical modelling has been a complementary tool in assessing possible patterning mechanisms. Until the last few years, these studies have focused on melanocytes and xanthophores, neglecting S-iridophores. The most commonly used mathematical paradigm for stripe formation takes the form of a Turing reaction-diffusion model. In these representations, melanocytes and xanthophores diffuse and interact via a few long- and short-range 'reactions'. This class of model typically rely on a small number of parameters which, upon being altered, can generate a diverse range of patterns. Minimal models such as these have the benefit that they are sometimes analytically tractable, allowing a deep understanding of the model. However, a potential limitation is that parameters do not always have a clear biological interpretation which, can sometimes make it difficult to

link parameters to measurable data. In the context of zebrafish stripe formation, these models have not yet incorporated S-iridophores (**Watanabe and Kondo, 2015**; **Kondo, 2017**; **Painter et al., 2015**; **Bloomfield et al., 2011**; **Binder and Simpson, 2013**; **Volkening and Sandstede, 2015**; **Kondo, 2017**; **Nakamasu et al., 2009**; **Moreira and Deutsch, 2005**; **Bullara and De Decker, 2015**; **Yamaguchi et al., 2007**; **Asai et al., 1999**). They suggest that the role for iridophores is restricted to simply orienting stripes (**Volkening and Sandstede, 2015**; **Nakamasu et al., 2009**; **Binder and Simpson, 2013**). New biological observations demonstrate that S-iridophores play a fundamental role in body stripe formation (**Singh and Nüsslein-Volhard, 2015**; **Frohnhöfer et al., 2013**; **Patterson and Parichy, 2013**). In particular, it has been shown that without S-iridophores, spots of melanocyte aggregates form instead of stripes, which is contrary to what these Turing reaction-diffusion models predict. These findings have paved the way for more detailed modelling, such as that of **Volkening and Sandstede, 2018**, who demonstrated (using an off-lattice individual-based model) the need for understanding S-iridophore behaviour when representing all three cell-types. For these reasons, we consider an inclusive modelling approach, incorporating the crucial cell-type S-iridophores and the full range of interactions depicted above.

Here, in a bottom-up approach, we hypothesise that the current biological understanding is sufficient to explain the major aspects of pigment pattern development and construct a model to test this. In particular, we construct an agent-based model incorporating all three pigment cell-types and their documented cellular interactions. We use observations of a set of three mutants that each lack an individual cell-type, plus the three double mutant combinations lacking pairs of cell-types, to deduce the key rules likely underpinning S-iridophore dynamics. Combining these assumptions with experimentally verified biological mechanisms in the literature, we generate a working model of adult pattern formation. We then run simulations for wild type (WT) and these mutant fish. We show that in each case our model correctly predicts the patterns observed in vivo, and that pattern development displays multiple quantitative matches to that in vivo using a parameter sampling methodology to demonstrate the robustness of these patterns to parameter variation. In an independent test of the model, we simulate mutants with pigment pattern defects caused by changes other than to the presence of pigment cell-types, and show that these too are successfully matched in silico by our model.

Our work demonstrates that current biological understanding, alongside simple assumptions about S-iridophore behaviour, is sufficient to explain adult pigment pattern formation in WT and multiple mutants. Our work reinforces the growing realisation in the field that the previously neglected S-iridophores are crucial for stripe formation, suggests a minimum set of their rules, and reveals unexpected subtleties to the phenotypic impact of the well-studied *leo* mutant.

Materials and methods

Modelling overview

We build our model with direct reference to the known biology. We model five cell types as individual agents: melanocytes (M), xanthophores (X), xanthoblasts (X^b), the unpigmented precursor cell to xanthophores) and S-iridophores in either dense or loose form (I^d, I^l, respectively). These are the cells we deem from the literature to be crucial for successful pattern formation. We do not directly model L-iridophores, since these appear after the adult pattern is formed and are more likely involved in pattern maintenance (**Frohnhöfer et al., 2013**). Unlike previous models of stripe formation (**Nakamasu et al., 2009**; **Bullara and De Decker, 2015**; **Volkening and Sandstede, 2015**; **Painter et al., 2015**; **Bloomfield et al., 2011**; **Volkening and Sandstede, 2018**), we include xanthoblasts as an independent cell-type in our model. This is because the larval xanthoblasts appear principally by dedifferentiation of the embryonic xanthophores, and most metamorphic xanthophores arise from the larval xanthoblasts (**Mahalwar et al., 2014**; **McMenamin et al., 2014**; **Budi et al., 2011**; **Singh et al., 2014**; **Dooley et al., 2013**), whilst xanthoblasts that do not re-differentiate into xanthophores persist in the stripe regions where they play a role in consolidating melanocytes into stripes.

Zebrafish pattern formation generates distinct pigment cell layers in the hypodermis (**Figure 1B**) a melanocyte, xanthophore and S-iridophore layer (**Hirata et al., 2003**; **Hirata et al., 2005**). For

consistency, we model each of the three layers as independent, two-dimensional lattice domains throughout pattern formation (***Figure 2A***).

Agents representing X and X^b, M, I^d and I^l occupy lattice sites, within xanthophore, melanocyte and S-iridophore domains respectively (***Figure 2A***). To account for the different packing densities of the cell types, lattice sites within the xanthophore and S-iridophore model layers are half the width and length of melanocyte sites size. This packing density does not have an impact on pattern formation, but, is included for biological realism. (For more details, see Appendix 1). Within each layer, volume exclusion properties hold: no two agents can occupy the same site at any one time (i.e. cells do not overlap).

The system is initialised to represent a typical WT fish shortly after the start of adult pigment pattern development ($\approx$25 dpf). We set the domain height to be 1 mm, since this is the approximate height of the fish at 25 dpf (***Supplementary file 3*** for details), we set the domain length to be 2 mm, representing approximately one-third of the full length, from the tip of the snout to the start of the tail, at 25 dpf, and thus equivalent to the trunk (***Parichy et al., 2009***). We populate the domain itself at $t = 0$ as an approximation of the observed larval pattern at 25 dpf (***Frohnhöfer et al., 2013***). At this time, there is a central stripe of dense S-iridophores along the horizontal myoseptum, scattered melanocytes and de-differentiated xanthophores (xanthoblasts) scattered across the domain.

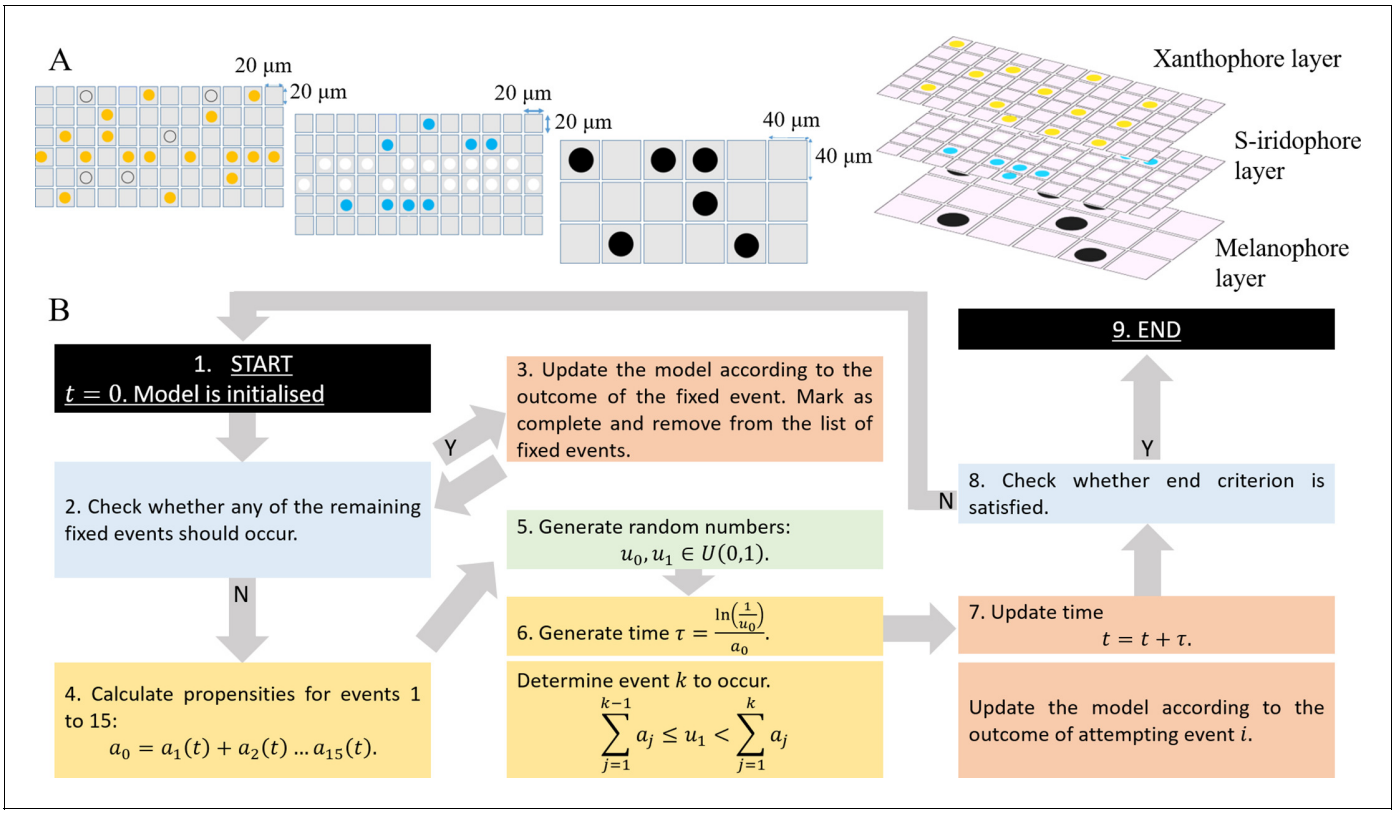


Figure 2. Model setup and simulation. (**A**) An example model setup. The domain is made up of three layers. Layer $\boldsymbol{X}$ which contains yellow X (yellow circles) and unpigmented X^b (clear circles with black outline). Layer $\boldsymbol{I}$ which contains silvery I^d (white circles) and blue I^l (blue circles). Layer $\boldsymbol{M}$ which contains black M (black circles) only. Each lattice site on each of the respective layers contain at most one cell at any given time. The layers are stacked on top of each other as seen in real fish. We note that in our simulations the ordering of the layers does not play any role in determining pattern formation. (**B**) Schematics of model implementation. (1) The model is initialised as described in Appendix 1. (2) The model is checked for the requirements for a fixed event to occur (e.g. new cells differentiating at a given time). If there is a fixed event to be implemented, the algorithm moves to stage 3. Otherwise, the algorithm passes to stage 4. (3) The model is updated according to the fixed event and algorithm returns to stage 2. (4) The propensity functions - probabilities for all other events to occur $\alpha_1(t), \ldots \alpha_{15}(t)$ are calculated. (5) Random numbers $u_0, u_1 \in U(0,1)$ are generated. (6) Numbers u_0, u_1 are used to determine the time τ until the next event and which event will be implemented. (7) The model configuration is updated. (8) The algorithm checks if the end criterion is satisfied, that is in the case of WT, that $\Omega_{SL} = 13.5mm$. If so, the algorithm finishes. Otherwise the algorithm continues, returning to stage 2. (9) The simulation completes.

We model this by populating the central three rows of the S-iridophore layer with dense S-iridophores, and by distributing melanocytes uniformly at random into sites within the melanocyte domain at density 0.04 and xanthoblasts uniformly at random into sites in the xanthophore domain at density 0.4.

The model is then updated according to the Gillespie algorithm (***Gillespie, 1977***). An overview of how the model is updated is given in ***Figure 2B*** and can be described as follows. At any given time t, the model is first assessed for meeting the criteria of a fixed event. Fixed events are all biologically determined events that occur once at a fixed time. For example at the start of pattern formation, the appearance of dense S-iridophores along the horizontal myoseptum is a fixed event. If the model meets the criteria, the fixed event occurs, is subsequently marked as complete and the simulation continues. If no fixed time event is to be implemented then one of 15 possible continuous time events is attempted. To do this, we treat all the potential actions, (for example cell birth or domain growth [as described in Section "Modelling assumptions"]), as individual 'events', each with an exponentially distributed waiting time which corresponds to their rate of occurrence (as specified in the literature ***Supplementary file 4***). To update the model at any given time $t = T$, an exponentially distributed waiting time; τ is generated until the next possible 'event' occurs (based on the rates of all of the possible events). Next, a random number $u_1 \in U(0,1)$ determines which event occurs based on the relative probability of each event occurring. Once an event is chosen, the domain is updated accordingly: if conditions required for that event to occur are met, the event is implemented, whereas if they are not then there is no change. Time is also now updated to $t = T + \tau$. This process repeats until we reach the end of pattern metamorphosis, defined by the simulated field standard length reaching approximately 13.5 mm (***Supplementary file 3***). The stochastic nature of our algorithm means that in any given simulation, the final pattern and its individual development will be inherently different to any other simulation, just as in real fish. Events incorporated into our model include all processes involved in the self-organisation of pigment cells during pattern metamorphosis as well as uniform domain growth with rate 0.13 mm per day in horizontal axis and 0.033 mm per day in the vertical axis (***Parichy et al., 2009***). These events are described in more detail in Section "Modelling assumptions".

Cells interact in the fish skin at both short (neighbouring cells) and long (up to half a stripe width $\approx$0.25 mm) range, with interactions thought to use direct contact through cellular extensions (filopodia, dendrites, or longer airenemes). In our model, uniform disks, with radii on the order of the distance between 2 cells ($\approx$0.04 mm) account for short-range interactions (***Figure 3A–D***), and an annulus with an outer radius of 0.24 mm (12 cells) and inner radius of 0.22 mm, (11 cells) represent long-range dynamics (***Figure 3E–H***). We allow cell interactions across different layers (as in real pattern formation). Cells that are chosen for movement can move into one of eight sites local to them. The probability of movement in one of the eight direction is biased according to how attracted or repelled the focal cell is to its local neighbours (***Figure 3I–J***). For more detail about how short- and long-range interactions are implemented see Appendix 1. See ***Supplementary files 4***, ***5*** and ***6*** for a detailed justification of the rates, interaction types and parameter values, respectively.

Modelling assumptions

In this section, we describe our modelling assumptions with regard to cell–cell interactions. These assumptions include all the known interactions between melanocytes, dense S-iridophores, loose S-iridophores, xanthophores and xanthoblasts, as well as some predictions about S-iridophore behaviour which have not been experimentally investigated in the literature. Apart from those involving S-iridophores, all the interactions and wherever possible their quantitative properties (strength, frequency etc) come directly from the literature, and are summarised in ***Figure 4G***, 1-14, and described in ***Supplementary file 5***. These include interactions influencing the movement, proliferation, differentiation and death of all cell types. These are represented explicitly in the model in as biologically realistic a manner as possible, at their determined rates.

The interactions involving S-iridophores have not been well-characterised experimentally, so we have developed our own predictions based on the literature describing S-iridophore behaviour during pattern metamorphosis. It has been shown using clonal cell analysis that during pattern metamorphosis S-iridophores spread across the skin of the zebrafish bidirectionally by proliferation of existing cells (between once and twice per day) combined with quick migration (***Mahalwar et al., 2014***). We further predict that dense S-iridophores show a directional bias towards xanthophores in

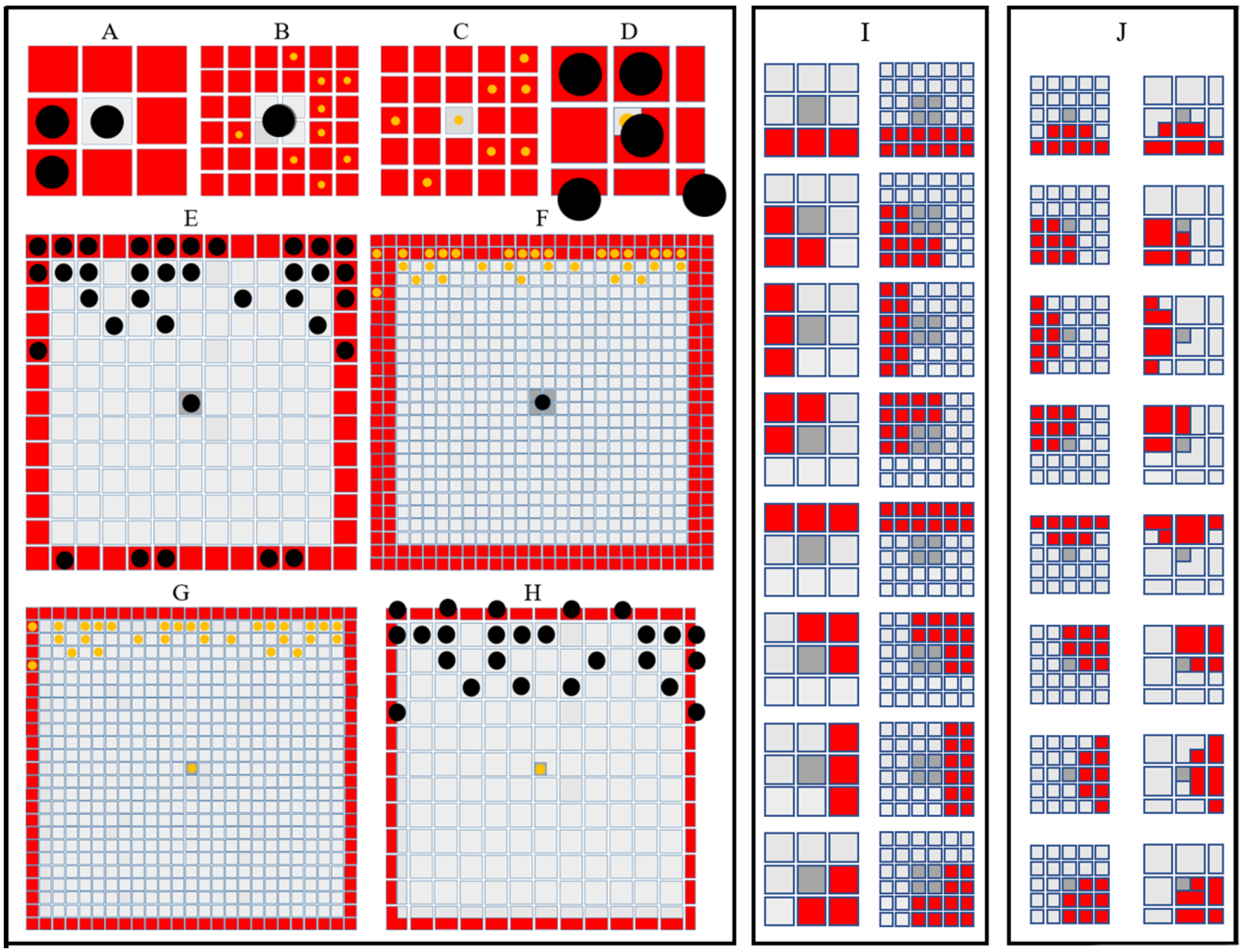


Figure 3. Simulating short- and long-range interactions. (**A–F**) Comparing the number of cells in the short (**A–D**) and long (**E–F**) range distance (0.04 mm) from a central site, on different domain types. M are represented as black circles. X are represented as yellow circles. (**A–B**) A visualisation of sites (marked in red) local to central melanocyte on the (**A**) melanocyte domain and (**B**) xanthophore domain.(**C–D**) A visualisation of sites (marked in red) local to central xanthophore on the (**C**) xanthophore domain and (**D**) melanocyte domain. (**E–F**) A visualisation of sites (marked in red) long-distance to central melanocyte on the (**E**) melanocyte domain and (**F**) xanthophore domain.(**G–H**) A visualisation of sites (marked in red) long-distance to central xanthophore on the (**G**) xanthophore domain and (**H**) melanocyte domain. (**I**) Melanophore movement with respect to local neighbours. If a melanophore located in a central position (marked in dark grey) attempts to move, the melanophore will consider neighbours in all sites marked in red on the melanophore domain (left), xanthophore and S-iridophore domain (right). (**J**) Movement on xanthophore/iridophore domain with respect to local neighbours. If a xanthophore/iridophore located in a central position (marked in dark grey) attempts to move, the xanthophore/iridophore will consider neighbours in all sites marked in red on the xanthophore and S-iridophore domain (left), melanophore domain (right).

the short range. We propose that dense S-iridophores are attracted in the short range to xanthophores since they are highly associated with each other in each of the mutants and in WT (***Frohnhöfer et al., 2013***), and this mutual attraction may be important for interstripe consolidation. Furthermore, loose S-iridophores show a directional bias away from other loose S-iridophores in the short range. We propose that loose S-iridophores are repelled by other loose S-iridophores as this would facilitate the prompt spreading of loose S-iridophores across the stripe regions.

Interestingly, as the S-iridophores spread they switch between a loose and dense form, predetermining the positioning of stripes and interstripes consecutively. In the loose form S-iridophores are spread and stellate in appearance. In contrast, in the dense form, S-iridophores are compact. The

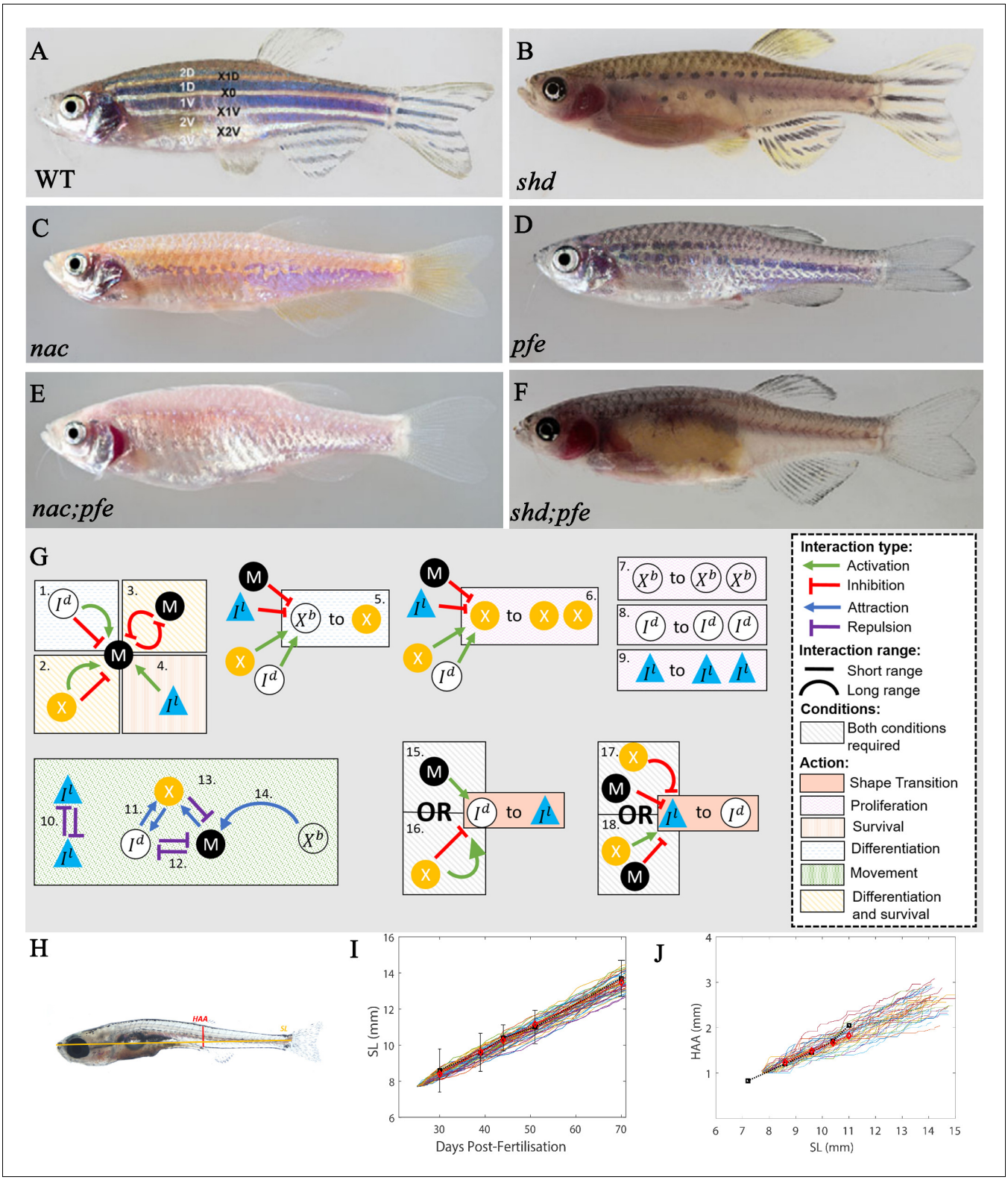


Figure 4. 'Missing cell' mutants and cell-cell interactions implemented. (**A–F**) Adult phenotype of the set of fish used to deduce S-iridophore interactions. (**A**) WT fish, (**B**) *shady* (*shd*), (**C**) *nacre*(*nac*), (**D**) *pfeffer* (*pfe*), (**E**) *nac;pfe*, (**F**) *shd;pfe*. (**G**) A representation of all the interactions implemented in the model. See ***Supplementary files 4*** and ***5*** for a detailed justification of the rates and interaction types, respectively. See ***Supplementary file 6*** for a detailed justification of all parameters. The direction of the arrow/inhibition sign indicates which cell is acting on which. For example in assumption

Figure 4 continued on next page

Figure 4 continued

13, the blue arrow from melanophore to xanthophore indicates that melanophores attract (green background) xanthophores. The purple inhibition arrow indicates that the xanthophores repel melanophores. OR statements in assumptions 15–18 are logical OR statements. For example, conditons 15 and 16 indicate that a dense S-iridophore may become loose if there are melanophores in the short range OR, there are no xanthophores in the short range AND many xanthophores in the long range. (**H**) A visualisation of measurements HAA and SL. (**I**) Plots of SL versus days post-fertilisation (dpf) for 40 model realisations. Each coloured line is a single realisation. Black squares are SL versus days post-fertilisation extracted from ***Parichy et al., 2009*** in real fish given in ***Supplementary file 3***. Red diamonds are the mean SL versus days post-fertilisation from 40 simulations. (**J**) Plots of SL (mm) versus HAA (mm) for 40 model realisations. Each coloured line is a single realisation. Black squares are SL versus HAA extracted from ***Parichy et al., 2009*** given in ***Supplementary file 3***. Red diamonds are the mean HAA versus SL from 40 simulations. Error bars are one standard deviation. The model agrees well with the data in both cases. Images (**A–F**) reproduced from ***Frohnhöfer et al., 2013*** and licensed under CC-BY 4.0 (https://creativecommons.org/licenses/by/4.0).

transition between the two types appears to be interchangeable. When dense S-iridophores initially spread beyond the boundary of the first X0 interstripe, they can later change to loose type (***Fadeev et al., 2015***). Similarly when loose S-iridophores reach an interstripe region, they can aggregate and form dense S-iridophores. It is not clear exactly what causes these shape transitions physically, and this is not a question we address here. It has, however, been shown that loss of Tjp1a function in *sbr* mutants compromises the transition of S-iridophores from dense to loose state, suggesting that Tjp1a contributes to regulation of the molecular switch that regulates S-iridophore shape changes during their dispersal (***Fadeev et al., 2015***). We envisage iridoblasts as initially differentiating in a dense form along the horizontal myoseptum, proliferate, migrate and spread, later de-differentiating and then re-differentiating into the opposite form dependent on their location with respective to other cell types (melanocytes and xanthophores).

Here, we hypothesise how the cell types affect S-iridophore type (loose or dense). The cause of these transitions is largely unknown, however, it has been suggested to be dependent on signals from melanocytes and xanthophores transmitted by gap junctions (***Irion et al., 2014***; ***Fadeev et al., 2015***). In order to investigate this, we consider a primary set of six mutants known to prevent the formation of one or more individual pigment cell-type. We use these to define the contribution and nature of S-iridophore interactions in pattern formation, by considering the outcomes in fish lacking each of the three cell types. Specifically we consider:

- Mutants lacking S-iridophores: The gene *shady* (*shd*) encodes zebrafish leukocyte tyrosine kinase (Ltk) which plays a role in S-iridophore specification (***Lopes et al., 2008***). As a result, strong *shd* mutants lack all S-iridophore types. The resultant adult pattern consists of a widened X0 region of xanthophores, which are flanked dorsally and ventrally by melanocytes organised as spot-like clusters in a sea of xanthophores, forming broken stripes (***Figure 4B***).
- Mutants lacking melanocytes: The gene *nacre* (*nac*) encodes the transcription factor Mitfa (***Lister et al., 1999***). *nac* mutants lack melanocytes throughout embryonic and larval development (***Lister et al., 1999***). As a result, stripes do not form properly and the adult phenotype consists of a prominent X0 interstripe of dense S-iridophores and xanthophores with irregular borders, accompanied by spots of dense S-iridophores and xanthophores ventrally. The rest of the flank is filled with loose form S-iridophores (***Figure 4C***).
- Mutants lacking the xanthophore lineage: Gene *pfeffer* (*pfe*) (alternatively known as *salz* (*sal*)) encodes colony stimulating factor one receptor (csf1ra) that is expressed and required specifically in xanthophores (***Frohnhöfer et al., 2013***; ***Patterson and Parichy, 2013***; ***Parichy et al., 2000b***). In the adult fish of strong alleles, xanthophores are almost absent in embryos, and absent in adults. The resultant adult pattern consists of a spotted melanocyte stripe pigmentation of normal alignment which fades out into a 'salt and pepper'-like pattern more posteriorly (i.e., in the tail) (***Figure 4D***). Melanocyte spots are associated with loose S-iridophores. In the regions lacking melanocyte aggregation (the 'salt-and-pepper' region), S-iridophores take a dense form, with melanocytes scattered at very low density, an arrangement never seen in WT patterns.
- Double mutants of the aforementioned mutant types: *nac;pfe*, *nac;shd* and *shd;pfe* (***Figure 4E,F*** depict the adult phenotypes of *nac;pfe* and *shd;pfe* respectively, there is no image available for *shd;pfe*). These mutants lack two of the aforementioned cell types. The resultant adult pattern is a uniform distribution of the remaining cell type (***Frohnhöfer et al., 2013***). These mutant phenotypes demonstrate that zebrafish stripe formation is not determined by an underlying pre-pattern, but instead is generated by cell-cell interaction.

Upon evaluating these mutants, we make the following deductions about S-iridophore shape transitions during pattern formation:

- S-iridophores are initially dense and cannot change shape autonomously. This is based on observations of mutants *nac;pfe* which only contain S-iridophores and in which the adult phenotype consists of dense S-iridophores in a coherent sheet across the domain (***Frohnhöfer et al., 2013***). In contrast, *pfe* and *nac* both exhibit loose and dense S-iridophores (***Frohnhöfer et al., 2013***), suggesting that both melanocytes and xanthophores are capable of facilitating S-iridophore shape transitions.
- Melanocytes in the short range promote the transition of dense to loose, conversely, a lack of melanocytes in the short range promotes the transition of S-iridophores from loose to dense. We propose these interactions for the following reasons. Firstly, in *pfe* and WT, dense S-iridophores are associated with lack of melanocytes, for example within the interstripes, whilst loose S-iridophores are associated with melanocytes, for example in the stripe region. Since we predict that melanocytes are required for dense S-iridophores transition the simplest assumption is that melanocytes promotes dense to loose transitions in the short range. Since loose S-iridophores can re-aggregate to dense form in *pfe* we assume that this signal is bidirectional and therefore a lack of melanocytes promotes the loose to dense transition.
- Xanthophores in the long range and lack of xanthophores in the short range promotes the transition from dense to loose; conversely, a lack of xanthophores in the long range as well as many xanthophores in the short range, promotes the transition from loose to dense. We propose these interactions for the following reasons. In *nac* and WT, dense S-iridophores are associated with xanthophores, whilst loose S-iridophores are associated with a lack of xanthophores (***Frohnhöfer et al., 2013***). Since S-iridophores initially appear in dense form and become loose for example in *nac*, when there are xanthophores in a low density local to S-iridophores and high density in the far range, we predict it is this combination that promotes the transition of dense to loose in the long range. Since in *nac*, S-iridophores can transition back from loose to dense when the local xanthophore density is high and far xanthophore density is low, we assume the opposing interaction is also true (***Frohnhöfer et al., 2013***).

These descriptions are summarised in ***Figure 4G***, 15-18. We note that in each of these cases variations of these interactions were already hypothesised by ***Frohnhöfer et al., 2013***. However, since their predictions did not distinguish loose and dense S-iridophores and did not indicate transition mechanisms, their predictions though similar, are extended here to incorporate these differences. The predictions we describe are the simplest possible for generating the patterns observed in the aforementioned set of fish, upon removing any one of these interactions, the model fails to generate the robust patterns we will describe (Figure 11 and Section "Necessity of S-iridophore assumptions").

Comparing simulated fish with real data

In order to validate our model, we compare different aspects of our simulation (size, spatial distributions of cells, numbers of melanocytes, stripe and interstripe width) with real fish at different developmental stages. In real fish, developmental stages are categorised according to the standard length (SL) of the fish (***Figure 4H***; ***Supplementary file 3***; ***Parichy et al., 2009***). For consistency, we calculate the 'stage' of our simulations using the length of our domain and a simple calculation to generate a simulated SL (see Appendix 1). This allows us to make a direct comparison between the range of sizes obtained in model simulations and the natural range in zebrafish HAA and SL. As a test of validity of this measure, ***Figure 4I*** and ***Figure 4J*** demonstrate 40 plots of simulated SL versus days post-fertilisation (dpf) and simulated HAA versus SL, respectively, compared with the averaged data (***Parichy et al., 2009***). These figures demonstrate that whilst growth rates are variable within simulations (as seen in real fish), the mean of our simulated rates approximately matches that in real fish.

Results

Modelling simulations

Having deduced this minimal rule-set from the literature and our further predictions from the phenotypes of the selected primary set of mutants, indicated in Section "Modelling overview" we use these as the basis for our model. We use our model to generate stochastic simulations of pigment pattern formation corresponding to the the period of adult metamorphic pigment pattern formation, during which the SL extends from 7.6 mm to 13.5 mm. We note that adult pattern metamorphosis and the appearance of metamorphic S-iridophores starts earlier, around 6–7 mm SL (***Parichy et al., 2009***). We initialise later at 7.6 mm as by this time the skin lying over the horizontal myoseptum is populated with an initial stripe of dense S-iridophores. We intialise our model accordingly to match this. Subsequently, we first assess the ability of the model to reproduce natural growth at a quantitative level, and then to generate the WT pigment pattern, both qualitatively and quantitatively. We then go on to simulate conditions corresponding to our primary set of mutants, considering the qualitative fit to the published patterns. To test robustness of the patterns, we provide a rigorous robustness analysis by carrying out one hundred repeats of the WT simulation and 'missing cell' mutants with perturbed parameter values chosen uniformly at random from the range 0.75–1.25 of their described value and show that in each case, the appropriate pattern is preserved. Finally, in a more rigorous test of the predictive power of our model, we explore three further mutant phenotypes that had not been considered in deriving the model's rule-set.

Simulation of WT pattern

In this section, we compare qualitatively our simulations of WT fish. For WT simulations, the model rules are given in Section "Modelling overview". ***Figure 5A–D*** depicts WT development, while ***Figure 5A'–D'*** (***Video 1***) shows a representative simulation using the model described by the rules in Section "Modelling overview". The simulations reproduce qualitatively most aspects of the biological pattern. The model is initialised at stage PB. At stage PR, we begin to see an accumulation of melanocytes either side of the initial interstripe and differentiation of new xanthophores. Furthermore, we see the development of 1D and 1V stripe regions and delamination of S-iridophores from dense to loose form at the edges of interstripe X0. At stage SP, we observe the spreading of loose S-iridophores across the two developing stripe regions. Finally at stage J+, we see three interstripes alternating with five dark stripes. The final pattern matches the stripes seen in the real WT fish and the cellular component of dark stripes (X, I^l, M) and light interstripes (I^d, X) matches the composition of pigment cells in real fish (***Figure 1C***). We emphasise that the simulations presented here (as well as in future sections) are a representative example of the model output.

Robustness of the model

Due to the abundance of parameters and cell-cell interactions necessary to capture what is known biologically about zebrafish pigment pattern formation, it is not feasible to perform an exhaustive parameter sweep to demonstrate the robustness of the model. Instead, as a test of robustness, we perform a rigorous robustness analysis by carrying out one hundred repeats of the WT simulation with perturbed parameter values chosen uniformly at random from the range 0.75–1.25 of their described value. The precise value of each parameter is sampled uniformly from this region, independentally for each parameter and each repeat. Twenty of these randomly sampled repeats are given in ***Figure 6***. We observe that for all one hundred repeats that small perturbations to the rates still generate consistent striping, demonstrating the robustness of the model.

Simulation of 'missing' cell mutants

In the next four sections (3.1.2–3.1.2), we compare qualitatively our simulations of mutants lacking one or more cell types. For the case of generating these mutants, we simulate the same WT model except we remove the appropriate cell type from the initial conditions and turn off cell birth of that cell type to match the mutation. For example, in *shd* we remove S-iridophores from the initial conditions and switch off S-iridophore birth. No other changes are made. For more information about mutant implementation see Appendix 2. These mutants often display similar features to WT fish; however, some aspects of the stripe formation are incomplete. In order to describe these

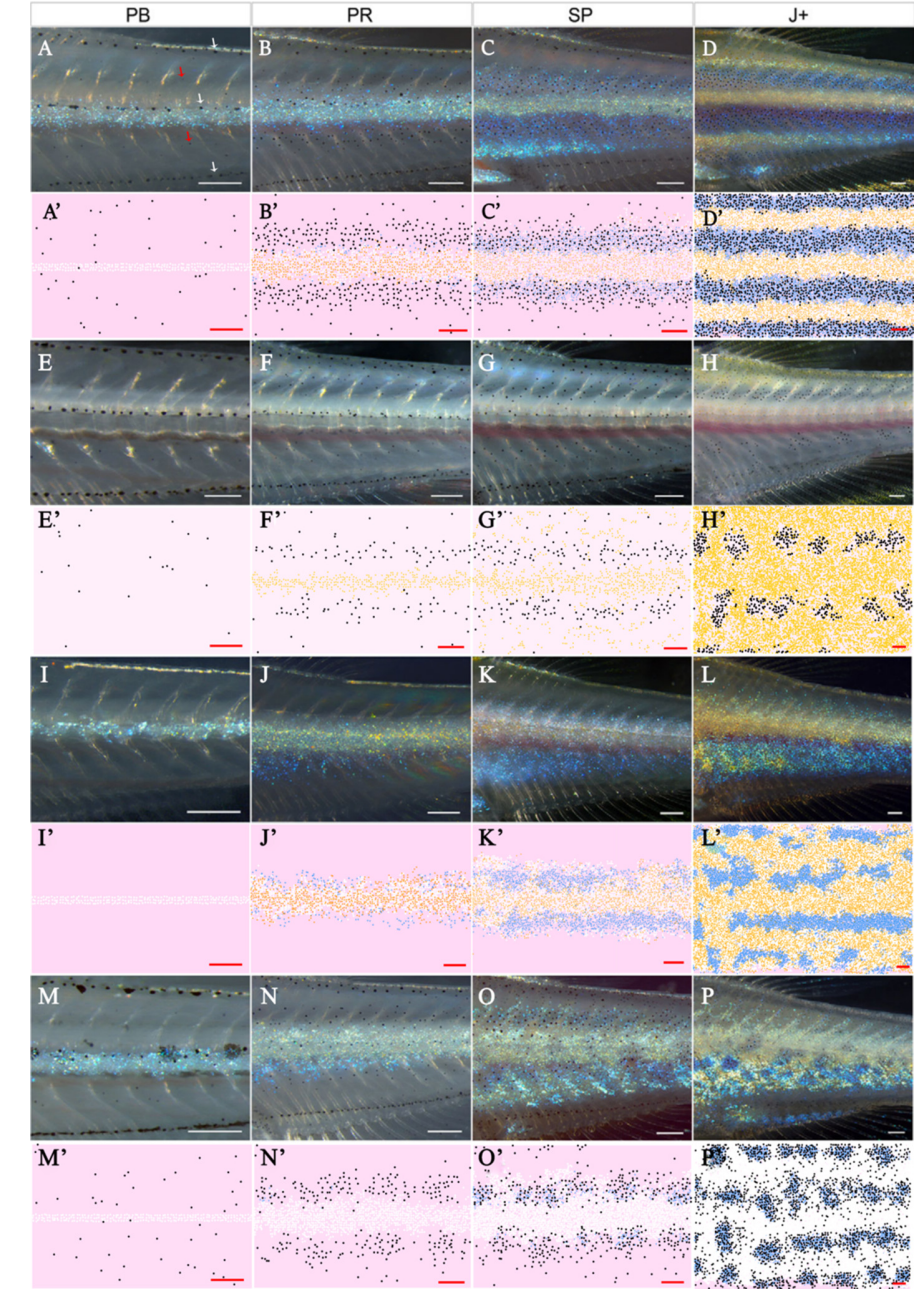
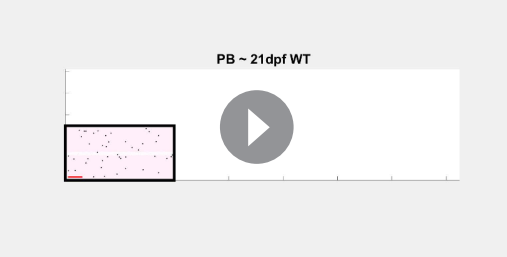


Figure 5. Representative simulation of WT, *shd*, *nac* and *pfe*. (**A–D**), (**E–H**), (**I–L**), (**M–P**) WT, *shd*, *nac*, *pfe* development respectively and (**A'–D'**), (**E'–H'**), (**I'–L'**), (**M'–P'**) corresponding model simulation. Red arrows: melanocytes in the skin layer, modelled in (**A**). White arrows indicate the embryonic pattern of melanocytes in four stripes, that are deeper than the skin level and are consequently not included in the model. See ***Videos 1–4*** for representative examples of these simulations in movie format. Scale bar is 0.25 mm for all images. Experimental images (**A–D**), (**E–H**), (**I–L**), (**M–P**) reproduced from ***Frohnhöfer et al., 2013*** and licensed under CC-BY 4.0 (https://creativecommons.org/licenses/by/4.0).



Video 1. Simulated development of WT.
https://elifesciences.org/articles/52998#video1

differences, with reference to *pfe*, *shd* and later in *sbr*, we define pseudo-stripes as the spots of melanophore aggregates that appear in a stripe-like orientation reminiscent of that in WT fish. We describe the pseudo-stripes in the order they appear as in WT fish. For example, we define pseudo-1D and pseudo-1V to be the first pseudo-stripes.

We demonstrate here the capability of our model simulations to reproduce qualitatively the pattern development of these mutants. For each mutant, we describe the initialisation of the model domain to match the fish at stage PB as well as all of the similarities observed between our model outputs and real fish at the following three developmental stages, PR, SP and J+. Finally, we describe the variation between many repeats of the model and how this correlates with real same-type siblings.

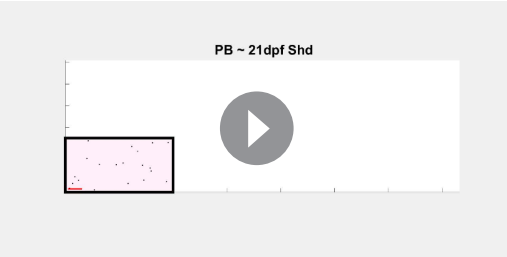
The *shady* mutant

At stage PB (***Figure 5E,E'***), we populate the domain with some randomly dispersed melanocytes at a lower density than that in WT (***Frohnhöfer et al., 2013***) and some randomly dispersed xanthoblasts at the same density as WT (***Frohnhöfer et al., 2013***). At stage PR (***Figure 5F,F'***), we observe some melanocytes beginning to differentiate in the usual 1D and 1V stripe regions. At stage SP (***Figure 5G,G'***), we observe the accumulation of melanocytes around the 1D and 1V stripe regions with a central stripe of xanthophores. Finally, at stage J+ (***Figure 5H,H'***), we observe two horizontal pseudo-stripes of melanocyte spots surrounded by xanthophores. We found that in 100 simulations, 100% of *shd* stage J+ mutants observed two pseudo-stripes (1D and 1V) just as in ***Figure 5H***. See ***Video 2*** for a movie of the simulation.

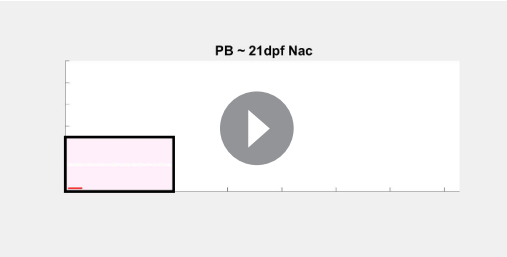
Moreover, pseudo-stripes varied in how stripe-like they were as observed in real fish (***Frohnhöfer et al., 2013***). As a measure of this, we calculated the longest stretch of melanophores in a row without any significant breaks over 100 simulations. This gives an indication of the widest 'spot' or 'pseudo-stripe' of melanophores in a simulation. We found that on the average, the mean of widest spot width over one hundred simulations was 0.18 of the simulated length. The widest spot width in 100 simulations was 0.43 of the simulated length, demonstrating the variance in pseudo-stripe length without break.

The *nacre* mutant

At stage PB (***Figure 5I,I'***) we populate the domain such that there is an initial stripe of dense S-iridophores and randomly dispersed xanthoblasts at the same density as in WT. At stage PR (see ***Figure 5J,J'***) we see the appearance of newly differentiated xanthophores associated with the dense S-iridophores in the initial X0 interstripe. At stage SP (***Figure 5K,K'***) we observe the switch of dense S-iridophores to loose form and the subsequent spreading of loose S-iridophores. Finally at stage J+ (***Figure 5L,L'***) we observe the jagged edges of

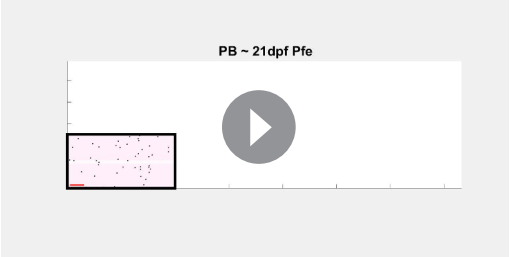


Video 2. Simulated development of *shd*.
https://elifesciences.org/articles/52998#video2



Video 3. Simulated development of *nac*.
https://elifesciences.org/articles/52998#video3

the usually straight X0 and the formation of a second pseudo interstripe some distance below X0 just as in real *nac* fish (***Figure 5L***). See ***Video 3*** for a movie of the simulation.



Video 4. Simulated development of *pfe*.
https://elifesciences.org/articles/52998#video4

The *pfeffer* mutant

At stage PB (***Figure 5M,M'***), we populate the domain with a central stripe of dense S-iridophores and randomly dispersed melanocytes at the same density as observed in WT (***Frohnhöfer et al., 2013***). At stage PR (***Figure 5N,N'***), we observe the arrival of melanocytes into the prospective 1D, 1V stripe regions that is less pronounced compared with WT simulations. At stage SP (***Figure 5O,O'***), we observe the accumulation of newly differentiated melanocytes into aggregates in prospective stripe regions 1D and 1V. Finally, at stage J+, (***Figure 5P,P'***) we observe the aggregation of melanocytes (associated with loose S-iridophores) into spots, surrounded by a sea of dense S-iridophores peppered with black melanocytes. In one hundred simulations, the median number of pseudo-stripes at stage J+ in these repeats was four, as WT. This is consistent with observations of *pfe* mutants, which typically show the same number of pseudo-stripes and -interstripes as WT fish (***Frohnhöfer et al., 2013***). We observe higher conservation of striping than in simulated *shd* mutants as observed in real fish. For example, in one hundred simulations, the average longest stretch of melanophores in any given simulation was 0.6 of the full length. See ***Video 4*** for a movie of the simulation.

As a test of robustness, we perform a rigorous robustness analysis by carrying out one hundred repeats of the mutant simulations with perturbed parameter values chosen uniformly at random from the range 0.75–1.25 of their described value as in Section "Simulation of WT pattern". Ten of these randomly sampled repeats are given in ***Appendix 4—figure 1***. We observe for all one hundred repeats and in all three mutants, that small perturbations to the rate parameters still generate consistent patterning, demonstrating the robustness of the model.

Double mutants; *shd;pfe*, *shd;nac*, *nac;pfe*

Lastly, we consider the double mutants. ***Figure 7A*** and ***Figure 7B*** depict adult patterns in *shd;pfe* and *nac;pfe* respectively. There is no image available for *shd;nac* adult or for the development of the aforementioned mutant phenotypes but it has been described in the literature that in all of the double mutants, the remaining cell type, by adulthood, fills the domain uniformly (***Frohnhöfer et al., 2013***). ***Figure 7A'–C'*** show a representative simulation for the mutants *shd;pfe*, *nac;pfe*, *shd;nac* respectively. In all cases of our model simulations, we observe that by stage J+ the remaining cell type begins to fill the domain. For example, in *nac;pfe* S-iridophores in dense form cover most of the flank by stage J+ (***Figure 7B'***).

Simulation of other mutants

In Section "Simulation of WT pattern", we demonstrated that our proposed model reproduces the WT, single and double mutant patterns and thus is sufficient to explain pattern formation in the skin. In this section, we perform a more stringent test of the model's completeness, by asking whether it can successfully simulate the outcomes of a set of pigment pattern mutants which were not used to deduce the rules underpinning our model. Since we were particularly interested to test our predictions of the rules relating to S-iridophore interactions, our secondary set comprises mutants with S-iridophore-related phenotypes: *rose* (*rse*) homozygotes, which show a reduction of S-iridophore numbers; *schachbrett* (*sbr*) homozygotes, which show a delay in S-iridophore shape transitions from dense to loose and *choker* (*cho*) homozygotes, in which the absence of the horizontal myoseptum prevents the formation of the initial dense X0 band of dense S-iridophores (***Figure 7D–F***). In the next few sections (3.1.3–3.1.3), we demonstrate the capability of our model to reproduce quantitatively the patterns of these mutants. In each section, we first describe the nature of the mutation and the way in which we adapt our WT model to simulate the mutants. We describe the similarities of our model simulation with real fish at the different developmental stages considered.

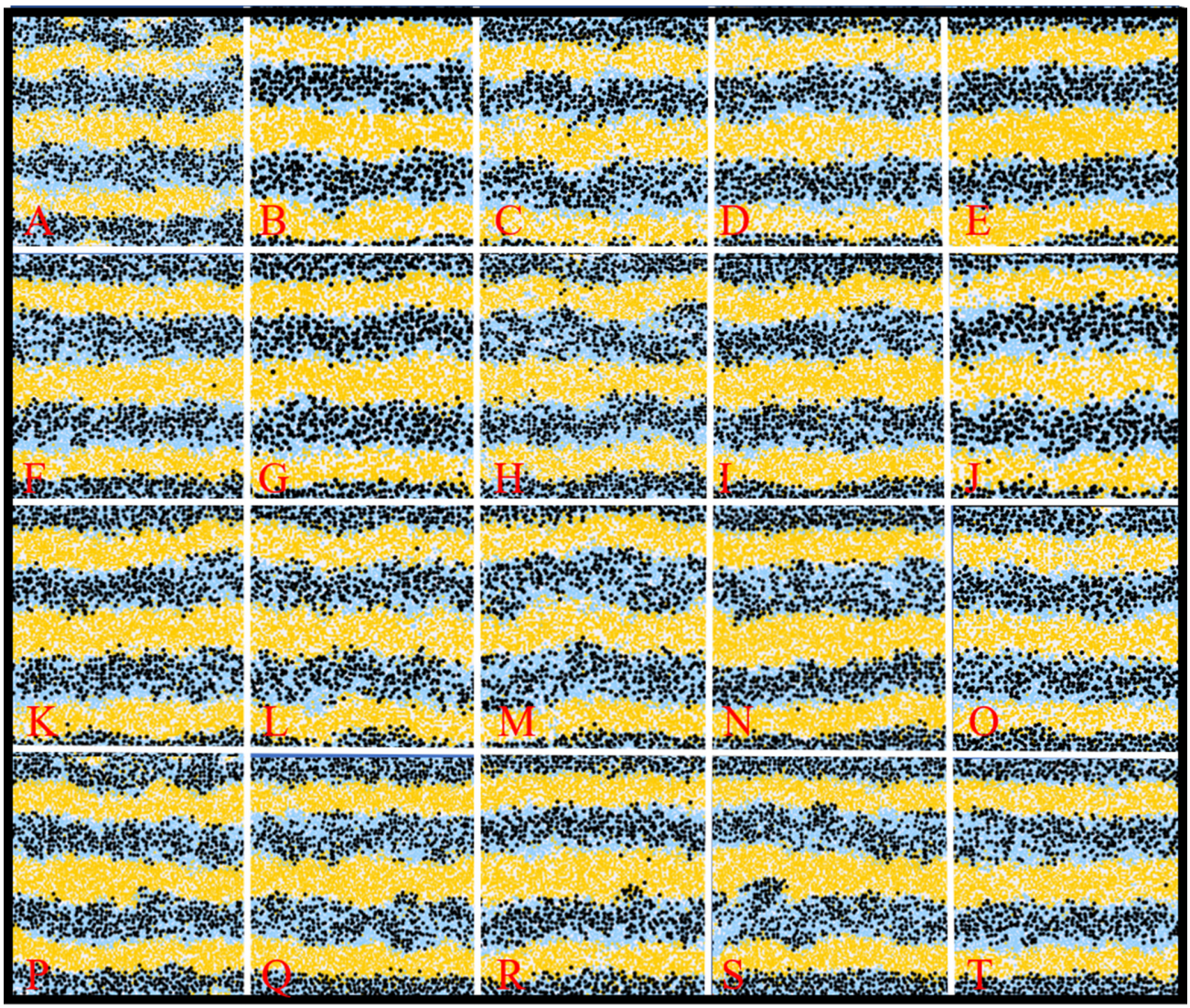


Figure 6. Example WT simulations at stage J+ when the parameters governing the rate of proliferation, movement, differentiation and death are perturbed slightly. Each square is an example WT simulation at stage J+ where each rate parameter is perturbed to $1+x$ times its normal value. The value x is chosen uniformly at random from the set $[-0.25, 0.25]$.

The *rse* mutant

Rose (*rse*), encodes the Endothelin receptor B1a (***Krauss et al., 2014***) and has been shown to acts cell-autonomously in S-iridophores; homozygous mutants result in a reduction of S-iridophores to approximately 20% of that seen in WT (observed in stage PB and adult fish [***Frohnhöfer et al., 2013***]). Consequently, adult fish show two broken dark stripes (reduction from four) bordering a widened X0 interstripe region. (***Figure 7D***). To simulate the *rse* mutant, we changed the number of initial dense S-iridophores at stage PB to one fifth of its usual number as observed in real fish at stage PB (***Figure 7G,G'***; ***Frohnhöfer et al., 2013***). At stage PR (***Figure 7H,H'***), there is a strong reduction in melanocyte number compared to WT (***Figure 7I,I'***) and we observe that dense S-iridophores spread less far from the horizontal myoseptum. At stage SP, (***Figure 7J,J'***) the stripe boundaries at X0 are poorly defined, and dense S-iridophores are still largely associated with the X0 region, with only a few loose S-iridophores appearing at the dorsal and ventral margins. At stage J+ (***Figure 7K,***

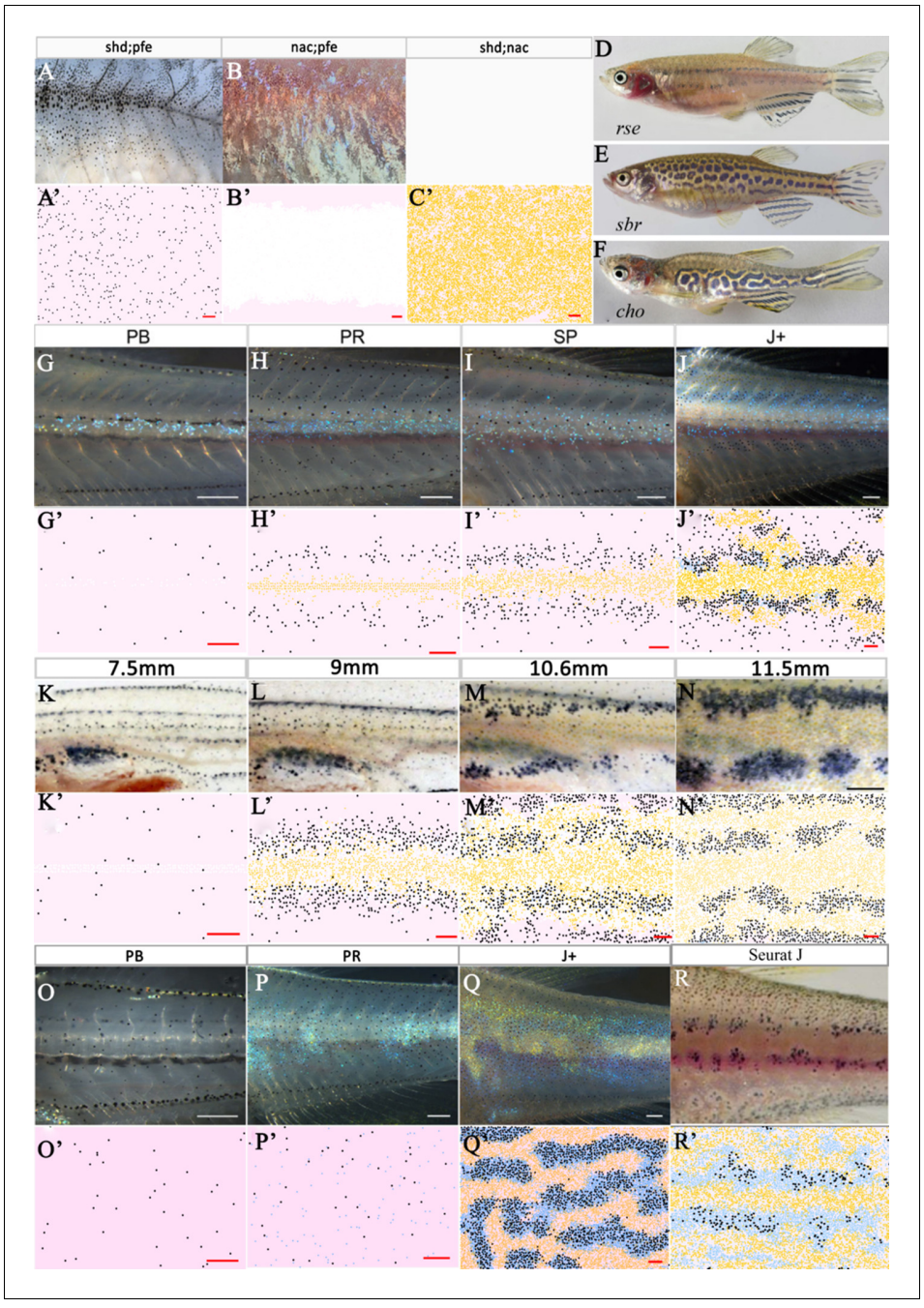


Figure 7. Other mutant fish simulations. (**A–C**) Real stage J+ *shd;pfe*, *nac;pfe*, *shd;nac* mutants, respectively. (**A'–C'**) Model simulations at stage J+ for *shd;pfe*, *nac;pfe*, *shd;nac* mutants, respectively. (**D–F**) Adult phenotype of selected (**D**) *rse*, (**E**) *sbr*, (**F**) *cho* mutant. (**G–J**), (**K–N**), (**O–Q**) *rse*, *sbr* and *cho* development, respectively. (**G'–J'**), (**K'–N'**), (**O'–Q'**) *rse*, *sbr* and *cho* model simulations, respectively. (**R**) *seurat* at stage J. (**R'**) seurat model simulation at stage J. Scale bar is 0.25 mm for all images. See ***Videos 5–7*** for representative examples of these simulation in movie format. Experimental images (**A–C**), (**D**), (**F**) (**G–J**) and (**O–Q**) are reproduced from ***Frohnhöfer et al., 2013***, (**E**), (**K–N**) from ***Fadeev et al., 2015***, (**R**) from ***Eom et al., 2012*** and are all licensed under CC-BY 4.0 (https://creativecommons.org/licenses/by/4.0).

Video 5. Simulated development of *rse*.
https://elifesciences.org/articles/52998#video5

K'), the stripe boundaries at X0 are more distinct, but the dark stripes are thinner and partly fragmented. See ***Video 5*** for a movie of the simulation.

The *sbr* mutant

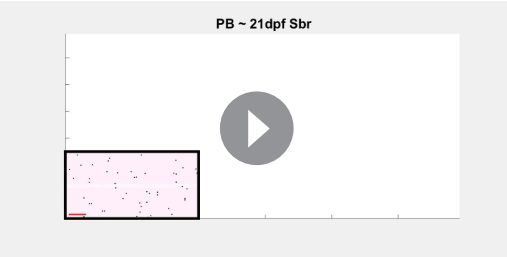
The *sbr* gene encodes Tight Junction Protein 1a (Tjp-1a), which is expressed cell autonomously in dense S-iridophores (but not in loose S-iridophores) and truncated in *sbr* mutants; in adult *sbr* mutants S-iridophore shape transition from dense to loose is delayed (***Fadeev et al., 2015***). As a result, adult fish exhibit interrupted dark stripes, generating a pattern reminiscent of a checkerboard (***Figure 7E***). ***Figure 7K–N*** depicts *sbr* early pattern development. During adult pigment pattern formation, differences from normal WT development are not seen until ≈10 mm SL (***Fadeev et al., 2015***), (SP stage) at which point instead of dense S-iridophores transitioning to loose S-iridophores at the edge of the interstripe, in *sbr*, dense S-iridophores remain dense and spread over melanocytes dorsally and ventrally bidirectionally . At later stages, some dense S-iridophores do switch to loose S-iridophores. See ***Video 6*** for a movie of the simulation.

We interpret the *sbr* mutation as causing a delay in signaling driving the transition of S-iridophores from dense to loose S-iridophore. We model this by reducing the attempted rate of transitioning from dense to loose to a rate 40 × less than the rate of attempting loose to dense S-iridophore transition. Due to available data, we initialise the model for *sbr* at 7.5 mm SL to match that published regarding the real fish (***Figure 7K,K'***). At 9 mm (***Figure 7L,L'***) melanocytes begin to accumulate either side of the widened initial X0 interstripe. At 10.6 mm SL, we observe melanocytes that are associated with dense S-iridophores (white cells) and not just with loose S-iridophores (blue cells) as usually seen in WT at 10.6 mm ≈ stage SA (between stages SP and J+, ***Figure 5C–D***). At 11.5 mm (***Figure 7M,M'***), melanocytes are organised into aggregates, approximately one stripe width in size, and only partially connected, thus forming a broken pseudo-stripe pattern.

The *cho* mutant

Homozygous *cho* mutant larvae lack the horizontal myoseptum (***Svetic et al., 2007***). As a result, dense S-iridophores are prevented from traveling via the horizontal myoseptum to generate the initial stripe of dense S-iridophores seen in WT at stage PB. Instead loose S-iridophores appear only later, at stage PR, uniformly across the domain. *cho* fish then proceed to develop a labyrinthine pigment pattern. Stripes and interstripes of normal width form in a parallel arrangement, but with orientation disrupted, with regions running vertically and horizontally and often strongly curved, sometimes branched and often interrupted (***Figure 7F***).

To model *cho*, we omitted the initial stripe of dense S-iridophores at the PB stage (***Figure 7O,O'***) and instead place 200 loose S-iridophores at random across the S-iridophore domain at stage PR (***Figure 7P,P'***). No other interactions were altered. At stage J+ (***Figure 7Q,Q'***), we see a pattern of normal width stripes and interstripes except with



Video 6. Simulated development of *sbr*.
https://elifesciences.org/articles/52998#video6

Video 7. Simulated development of *cho*.
https://elifesciences.org/articles/52998#video7

varying orientation, as seen in real *cho* fish. See ***Video 7*** for a movie of the simulation.

The *seurat* mutant

Homozygous *seurat* mutants develop fewer adult melanophores, thus forming irregular spots rather than stripes. This phenotype arises from lesions in the gene encoding Immunoglobulin superfamily member 11 (Igsf11) (***Eom et al., 2012***) which encodes a cell surface receptor (containing two immunoglobulin-like domains) that is expressed autonomously by the melanophore lineage. Igsfl1 promotes the migration and survival of these cells during adult stripe development as well as mediating adhesive interactions in vitro.

To model *seurat*, we reduced the rate at which melanocytes could differentiate to a twentieth of the usual rate. This was to reflect the inhibition of the migration of melanoblasts (precursors of melanophores) across the domain and increased the rate of attempted melanocyte death to one hundred times per day (usually once per day). No other interactions were altered. At stage J (***Figure 7R,R'***), we see a pattern of normal width stripes broken into spots with a reduced number of melanocytes. Melanocytes are associated with loose S-iridophores and xanthophores with dense S-iridophores, as seen in real *seurat* fish.

By modelling *seurat* and *sbr*, we can also make predictions about the phenotype of a double mutant *seurat;sbr*, shown in ***Appendix 5—figure 1***. We predict that by stage J+, this mutant would be covered in dense S-iridophore and associated xanthophores, with a few melanocytes at the very dorsal and ventral region of the fish. We are not aware of a published description of the phenotype of this double mutant, so our prediction remains to be tested.

Regeneration experiments

In order to further validate our model, we test to see whether we observe similar behaviour when the cells are ablated and the pattern is left to regenerate. In 2007, ***Yamaguchi et al., 2007*** ablated a rectangular window of pigment cells of adult zebrafish stripes and recorded the regeneration of pigment producing cells (***Figure 8A–C***). They found that after ablation, cells regenerated in a labyrinthine pattern. To model this ablation, we simulated WT development from stage PB to our latest simulation stage, J+ as seen in ***Figure 5D'***. At stage J+, we simulate ablation by removing a square region of cells in the centre (horizontally) of three stripes and interstripes. We then observe the pattern regeneration 7, 14 and 21 days later in ***Figure 8C–E***. At 14 days, we observe the production of irregular shaped spots of melanophores in the centre of the ablated region as seen in the ablated fish at day 7. At day 21, we observe a regeneration of the pattern where stripes are no longer oriented horizontally. In some regions, spots of melanophores are surrounded by xanthophores.

In 2013, ***Patterson and Parichy, 2013*** ablated a section of dense S-iridophores along the horizontal myoseptum using a S-iridophore-specific marker *pnp4a*:NTR+Mtz at the beginning of pattern metamorphosis (***Figure 9A***, stage PB). They then observed the subsequent pattern formation (***Figure 9B***). We simulate this by removing a section of dense S-iridophores from the horizontal myoseptum at stage PB (***Figure 9C***) and then simulating as normal. We observe the pattern at 10 days post ablation. Xanthophores are associated with the undamaged portion of the S-iridophore stripe and melanophores surround the damaged interstripe (***Figure 9D***). In both cases, our simulations closely approximate the published observations.

Quantitative analysis of simulations

In the next few sections (3.2.1)-(3.2.4), we test the consistency of our WT and mutant simulations, averaged over 100 simulations, with real published quantitative measures. We test four criteria using experimental data: (1) the number of melanophores in mutants at different stages with respect to WT at the same stage, (2) the average width of X0 interstripe for WT, *rse*, *pfe*, *shd* and *nac* at stage J+, (3) WT stripe straightness, (4) WT pigment cell density in stripes and interstripes.

Melanophore density across WT and mutant fish

Figure 10A is a comparison between the average number of melanocytes per ventral hemisegment for each mutant with the number of melanocytes at WT at stages PB, PR, SP, J and J+ using the data from ***Frohnhöfer et al., 2013***. First, since we do not have simulated data for a whole hemi-segment we normalise our melanocyte numbers against WT numbers at each stage. We do this by, for each respective stage and each respective mutant, multiplying the number of simulated melanocytes

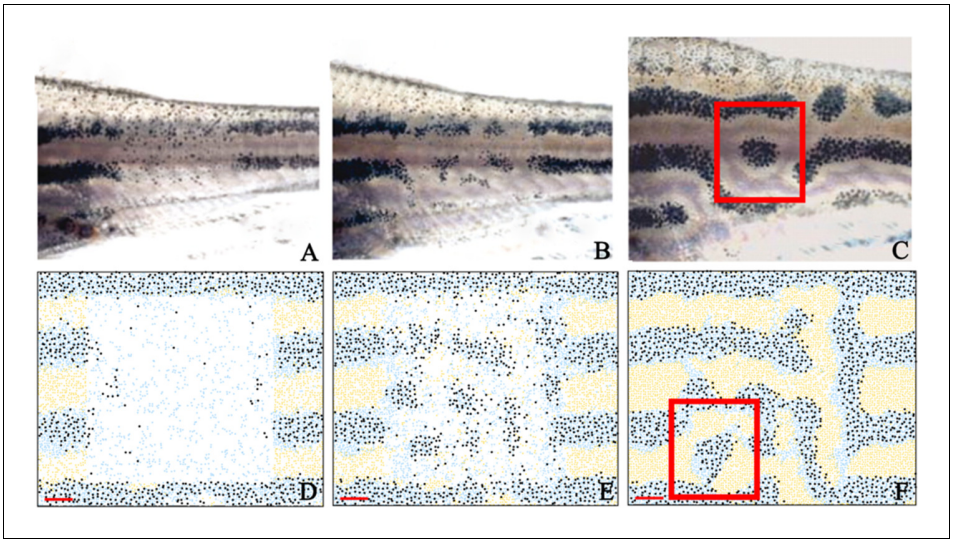


Figure 8. Stripe regeneration simulations. (**A–C**) Regeneration of new pigment producing cells 7, 14 and 21 days respectively after a small rectangular window of cells in the adult WT stripes are completely ablated (***Patterson and Parichy, 2013***). (**D–F**) A representative simulation of the regeneration of an adult zebrafish 7, 14 and 21 days after a simulated ablation has occurred. Scale bar is 0.25 mm in all images. Experimental images (**A–C**) are reproduced from ***Yamaguchi et al., 2007***. Copyright (2007) National Academy of Sciences U.S.A.; these images not covered by the CC-BY 4.0 licence and further reproduction of this panel would need permission from the copyright holder.

at the given stage for the given mutant by the number of melanocytes at each stage in real WT fish and dividing by the number of melanocytes observed in our WT simulations at this stage. These comparisons are given in ***Figure 10A***. We observe the same trends seen in real fish. Moreover, in all stages, except for stage SP, the simulated data falls within the error bars of the measured data in real fish. In particular, we observe that the number of melanocytes remains similar to WT in *pfe* until stage J+, similar to that in WT, whilst the number of melanocytes is significantly lower in *shd* and *rse* in comparison to WT just as in real fish.

WT stripe straightness

In real life, zebrafish stripes are quite straight, but are not necessarily perfectly so (see X0 in ***Figure 1A*** for example). To measure stripe straightness, we first generated a line representation (x) of the central interstripe (see Appendix 3). From this line x, we calculated the stripe straightness SS (x), measured by the ratio of the length of our line (L) to the straight line distance between the ends of it (C), *i.e.*

$$\mathrm{SS}(x) = \frac{C}{L}. \tag{1}$$

The value of SS(x) lies between 0 and 1, since $C \leq L$. For more information about how $SS(x)$ is generated see Appendix 3. In 10 real WT stage J+ fish, the mean SS value was 0.98. In 100 simulations we observe a high albeit slightly lower mean SS value of 0.92 at stage J+, demonstrating good stripe straightness, that is close to that observed in real fish.

X0 interstripe width across WT and mutant fish

Finally, we compare the interstripe X0 width in our simulations with real data. We choose this interstripe as it is the only interstripe that the mutants we consider and WT have in common. We

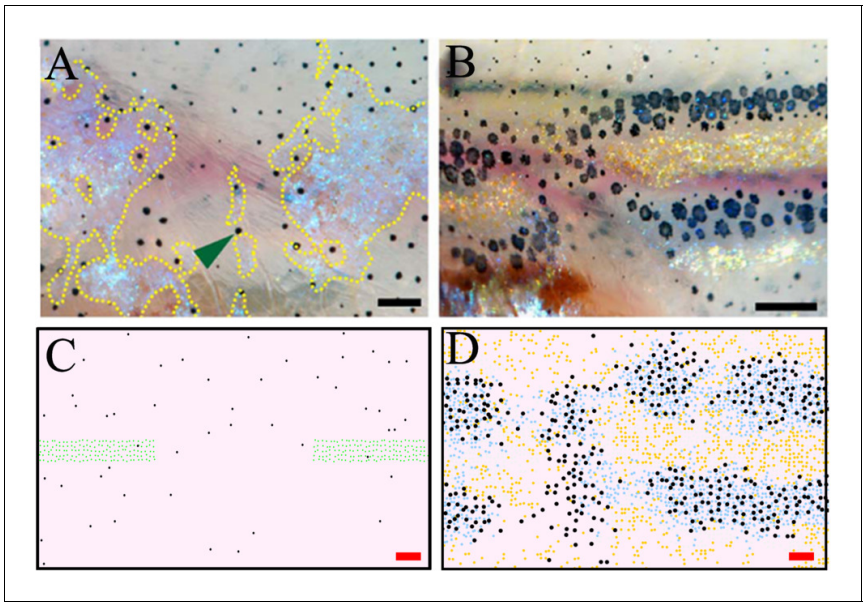


Figure 9. Stripe regeneration post S-iridophore ablation. (**A**) S-iridophores are ablated using *pnp4a*:NTR+Mtz at stage PB (**B**) Stripe development in wild-type fish following S-iridophore ablation. Xanthophores are associated with S-iridophore interstripe. Melanophores aggregate around the new interstripe. (**C–D**) A representative simulation of S-iridophore ablation and subsequent development. For clarity, dense S-iridophores are displayed in green in (**C**). Scale bar is 0.25 mm in all images. Experimental images (**A–B**) are reproduced from ***Patterson and Parichy, 2013*** and licensed under CC-BY 4.0 (http://creativecommons.org/licenses/by/4.0).

compare the width of this interstripe at J+ in our simulations (see Appendix 3 for detailed methods) with the observations of the corresponding J+ mutant in ***Frohnhöfer et al., 2013***. We demonstrate in ***Figure 5B*** that, in all cases, the real J+ mutant X0 interstripe widths fall in the range of ±1 standard deviations of the simulated J+ X0 stripe width averaged over 100 simulations. This demonstrates consistency between our model and real data. Thus, our model demonstrates an excellent ability to quantitatively simulate the patterns of real fish.

Pigment cell density in WT

There are no published estimates of WT pigment cell density in each of the stripes at the J+ stage when our simulations end. However, our data are comparable to those of adult WT fish measured by ***Mahalwar et al., 2016***, who observed that in the stripe regions there were approximately four times more xanthophores in the interstripe region than melanocytes in the stripe region. Furthermore, whilst the light interstripe were completely devoid of melanocytes, there was a low density of xanthophores in the stripe region. In our model simulations, we observe a mean of 4.01 times as many xanthophores in the interstripes than melanocytes in the stripes demonstrating good agreement. We also observe a low density of xanthophores in the stripe regions and negligible numbers of melanocytes in the interstripe regions.

Simulation reproducibility of pattern formation

To further test the accuracy of our model's outputs, we compare the spatial correlation of different cell types at different distances. We use this measure as an objective test of whether the spatial distributions between cells we observe in our representative simulations, (i.e. the patterns generated), are consistent among different simulated outputs.

To measure spatial correlation, we use a pair correlation function (PCF). A PCF determines whether, given a spatial distribution of agents on a domain, the number of pairs of agents at a certain distance from each other are greater than or fewer than the number expected if the agents were distributed uniformly at random. For example, if the PCF value is unity for a certain distance, this indicates that there is no spatial correlation. If the PCF value is greater or less than unity for a

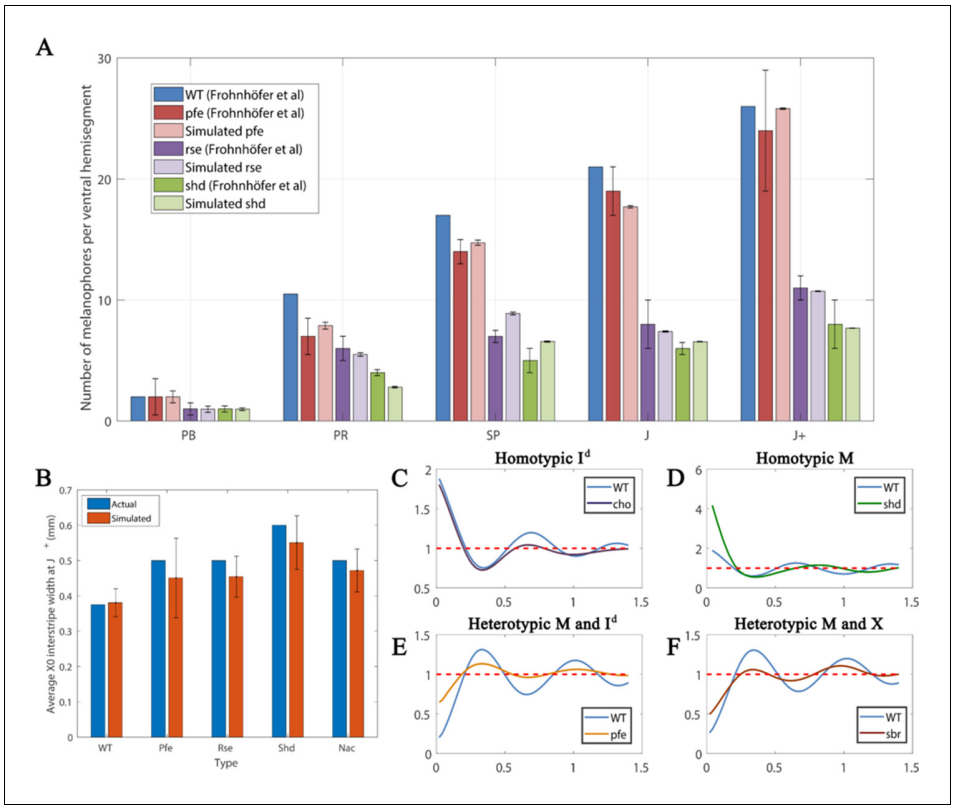


Figure 10. Quantitative analyses. (**A**) Number of melanocytes per ventral hemisegment. Quantification of melanocytes. A comparison of melanocyte numbers between mutants and WT seen in our simulations and data recreated from ***Frohnhöfer et al., 2013***. For each stage, our simulated number of melanocytes was normalised by the WT number in data by ***Frohnhöfer et al., 2013***. (The number of animals used for counting was at least five for each measurement point.) The number of simulations was 100. Error bars indicate the standard deviation. There is no error bar for WT, since WT data was used to normalise the data. (**B**) X0 interstripe width. A comparison of X0 stripe width in 100 simulations with X0 stripe width in real mutants (taken from representative images; ***Frohnhöfer et al., 2013***). The darker bars depict experimental data, lighter bars depict our simulated data. Error bars indicate the standard deviation. Since there is only one image for each of the respective mutants, there is no error bar for actual (real experimental data). (**C–F**) PCF for different cell types averaged over 100 simulations using the square uniform PCF (***Gavagnin et al., 2018***). The spatial pair correlation as a function of distance for (**C**) dense S-iridophores only (**D**) melanocytes only, (**E**) melanocytes and dense S-iridophores, (**F**) melanocytes and xanthophores. Experimental data in (**A**) is reproduced from ***Frohnhöfer et al., 2013***.

certain distance then this indicates that there is spatial correlation or anticorrelation respectively at that distance. The PCF we employ is specific to on-lattice domains and is called the Square Uniform PCF (***Gavagnin et al., 2018***) adapted for multiple cell-types (see Appendix 3 and ***Dini et al., 2018***). We describe the PCF as homotypic when we are measuring the spatial correlation of one cell-type and heterotypic when we consider the spatial correlation between two different cell-types. We choose this PCF as it uses a measure of distance that is complementary to the distance measurement in our simulations.

Figure 10C–F shows the square uniform PCF against distance for different mutants and cell types averaged over one hundred simulations. For each plot of a given PCF type (homotypic melanocytes, for example), we repeatedly simulate the relevant mutant to its final stage, compute the PCF of the resultant pattern and then average the PCFs over the number of repeats to give us an averaged PCF.

To interpret the data, consider the representative simulation for a WT fish at stage J+. In this example, we observe stripes (***Figure 5D'***). This is a consistent feature for all of the repeat simulations of WT at stage J+. To quantify average interstripe width (the distance vertically in mm from the top

of any interstripe to the bottom) in our simulations we can consider the averaged homotypic dense S-iridophores PCF for WT in ***Figure 10C***. We observe that this shows periodicity (sequential peaks and troughs) at different distances. These are a consequence of the striped pattern at J+ (***Figure 5D'***). Since dense S-iridophores occupy the interstripe regions in WT and not the stripe regions at J+, dense S-iridophores are spatially correlated at short distances, indicated by a positive value of the PCF at short distances. Conversely, they are anti-correlated at distances approximately one half, one and a half or two and a half stripe widths away, as these distances correspond to the relative positions of the dark stripes, which normally lack dense S-iridophores. We see troughs at these distances. The period of the PCF in this case thus quantifies an estimate for average interstripe width.

In the next few paragraphs, we test the reproducibility of different features that are observed in our representative simulations by considering a PCF of appropriate cell types averaged over one hundred repeats.

Real *cho* mutants and WT fish share similar interstripe width (***Frohnhöfer et al., 2013***)(also seen in our model; compare ***Figure 7Q'*** and ***Figure 5D'***). To test reproducibility we consider the homotypic PCF of dense S-iridophores for WT and *cho* in ***Figure 10C***. For both *cho* and WT simulations, the averaged homotypic PCF for dense S-iridophores observes periodic behaviour with the same frequency, indicating maintenance of interstripe width between WT and *cho* in our model, consistent with observations in vivo.

In real *shd* at stage J+, there are two pseudo-stripes of melanocytes broken into spots, of a diameter that is approximately equal to the normal stripe width (***Frohnhöfer et al., 2013***) (also seen in our model; compare ***Figure 5H'*** and ***Figure 5D'***). To test whether this is consistent we consider the homotypic PCF of melanocytes for WT and *shd* in ***Figure 10D***. The average stripe width of WT and the average aggregate size of *shd* can be approximated from the PCF as the distance related to the first trough, as this is the shortest distance at which melanocytes are most anticorrelated with other melanocytes. For both *shd* mutants and WT simulations, these are both approximately 0.3 mm.

In real *pfe* at stage J+, stripes and interstripes remain aligned and have the same width as in WT, except that stripes are broken into spots and some melanocytes lie ectopically in the usual interstripe region (***Frohnhöfer et al., 2013***) (also seen in our model; compare ***Figure 5D'*** and ***Figure 5P'***). To test reproducibility, we consider the heterotypic PCF of melanocytes and dense S-iridophores for WT and *pfe* in ***Figure 10E***. For both the WT and *pfe* simulations, the averaged heterotypic PCF of melanocytes and dense S-iridophores displays periodic behaviour with the same period. However, in *pfe* the peaks and troughs are damped. We interpret this as follows. Firstly, this indicates that, in our model, stripe width is preserved between *pfe* and WT as the period of the PCF is the same. Moreover, as the peaks and troughs are damped in *pfe*, this indicates that, as seen in our representative simulation, some melanocytes tend to remain in the interstripe regions.

In real *sbr* at 11.5 mm SL, stripes are sometimes broken into spots of usual (vertical) width, but the overarching stripe pattern remains (***Fadeev et al., 2015***) (also seen in our model, compare ***Figure 7D'*** and ***Figure 5N'***). To test reproducibility, we consider the heterotypic PCF of melanocytes and xanthophores for WT and *sbr* in ***Figure 10F***. The first peak of this PCF corresponds to the shortest distance at which melanocytes and xanthophores are most correlated, which is approximately the stripe width. For both *sbr* mutants and WT simulations these are both approximately 0.3 mm.

In these examples, using appropriate PCFs we have demonstrated the consistency of our simulations in generating patterns that match the qualitative differences we expect when we compare mutant fish with WT. We note that we have only provided the averaged PCF for the scenarios aforementioned for simplicity. For information about how the PCF is calculated, please see Appendix 3.

Iridophore assumptions and biological redundancy

Necessity of S-iridophore assumptions

For the less well-studied S-iridophore transitions, we analysed key mutant phenotypes to infer biologically realistic rules for these interactions, aiming to generate assumptions that were the simplest for pattern formation changes seen, but no simpler. These deductions are discussed in Section "Modelling overview". In ***Figure 11*** B1–J3, we demonstrate the necessity of all of the assumptions for dense-to-loose, loose-to-dense S-iridophore transitions first outlined in ***Figure 4G***, 15-18 for stripe formation by showing representative images (the model was run 50 times each) of simulations

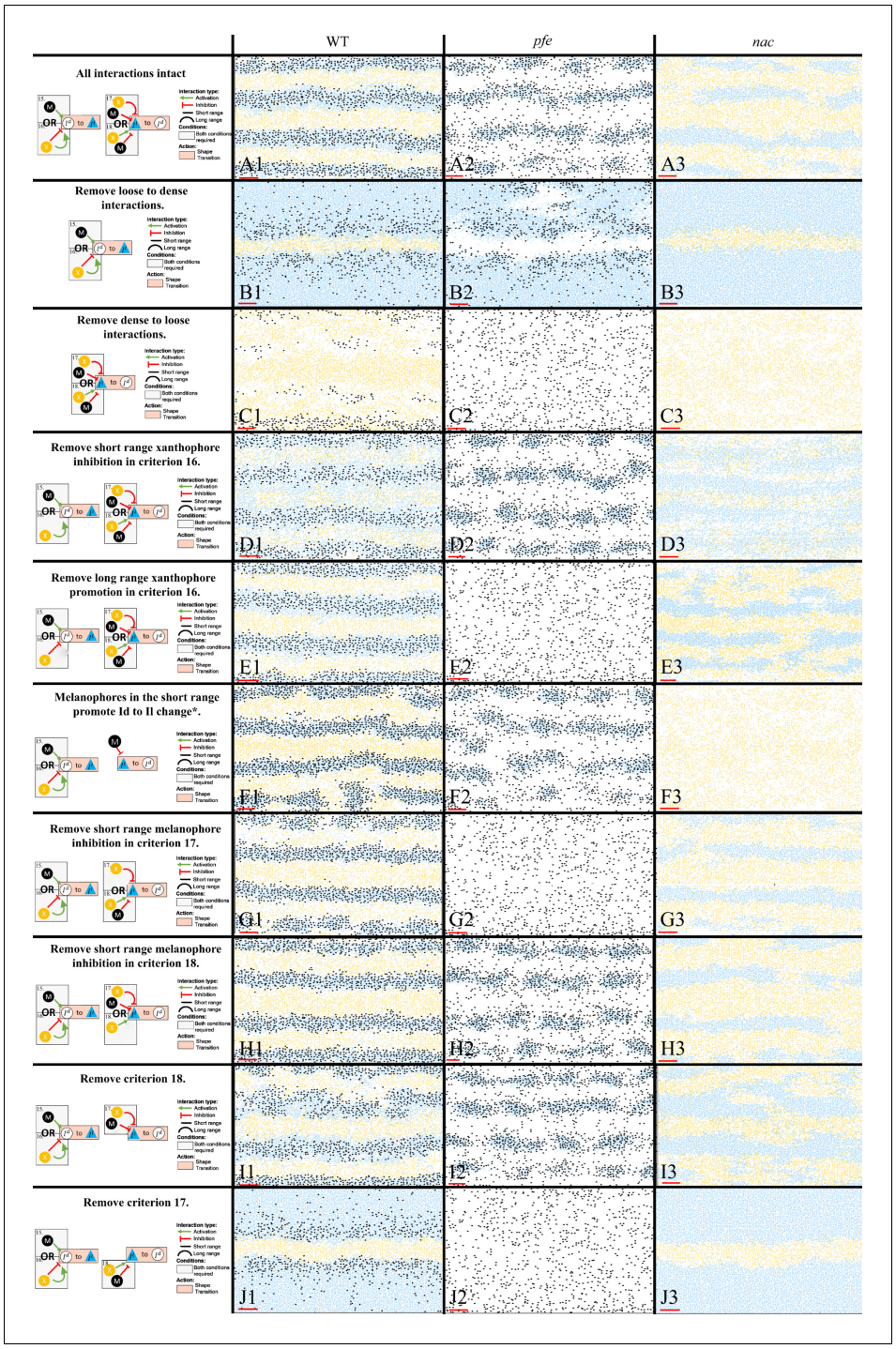


Figure 11. Representative simulations of the model with some of the S-iridophore interactions removed. The first column displays a diagram of the S-iridophore interactions which remain (all other cell-cell interactions are unchanged). Columns 2–4 are representative simulations of WT, *pfe* and *nac* under these conditions. *This interaction is equivalent to removal of long range xanthophore inhibition in criterion 17. It is also equivalent to removal of short range xanthophore promotion in criterion 18.

at stage J+ for fish lacking one or more S-iridophore transition mechanisms display patterns which diverge from those seen in real fish.

First, we analyse simulations lacking one of the transition types loose-to-dense or dense-to-loose (***Figure 11*** B1–C3). Without a loose-to-dense transition (***Figure 11*** B1–B3), note how in all cases (WT, *pfe*, *nac*) only one pseudo-interstripe is preserved: the initial X0 interstripe. The X0 interstripe is surrounded by a sea of loose S-iridophores which have transitioned to loose either by promotion by xanthophores in the long range (*nac* ***Figure 11*** B3), promotion by melanophores in the short range (*pfe* ***Figure 11*** B2) or a combination of both (WT ***Figure 11*** B1). This suggests that loose-to-dense interactions are important for generating subsequent interstripes. Interestingly, without loose-to-dense transitions WT fish demonstrate a striped pattern similar to *Danio albolineatus* (***Parichy, 2006***) suggesting a possible route of evolution between these fishes (also noted by ***Volkening and Sandstede, 2018***). Without a dense-to-loose transition (***Figure 11*** C1–C3), the S-iridophores form a dense sheet over the entire domain with xanthophores and melanophores scattered across the domain. This demonstrates the necessity for S-iridophores to be able to transition between dense and loose form.

Next, we consider removing each of the criteria required for an S-iridophore transition one at a time (***Figure 11*** D1–J3). We first note that, in some cases, removal of an interaction in either *nac* or *pfe* results in loss of a transition type. These are not shown in ***Figure 11*** for simplicity. In other scenarios, however, removal of an interaction leads to the uninhibited possibility of a transition in one of the single cell mutants. For example, consider ***Figure 11*** E2. Removal of long-range xanthophore promotion in criterion 16, leads to the possibility of a transition from loose to dense provided that there are either melanophores in the short range *or* no xanthophores in the short range. Since in *pfe* there are are always no xanthophores in the short range, or indeed anywhere on the domain, S-iridophores are consistently promoted to dense type, with a non-zero rate, thus ***Figure 11*** E2 is not distinguishable from ***Figure 11*** C2 since in effect, in *pfe* the same interactions have been knocked out. We note that this is the case for one of either *nac* or *pfe* in ***Figure 11C,E–G,J***.

In ***Figure 11*** D1–D3, short-range xanthophore inhibition is removed from criteria 16. As this is exclusively a xanthophore-iridophore interaction this only effects WT and *nac*. Without the promotion of xanthophores in the short range, S-iridophores change from dense to loose in the interstripes, making the interstripes appear faded. In ***Figure 11*** G1–G3, short-range melanophore inhibition is removed from criteria 18. Therefore S-iridophore can change from loose to dense when there are xanthophores in the short range. As a result interstripes become wider as they are unrestricted by local melanophore stripes. Finally, in ***Figure 11*** I1–I3 criteria 18 is removed, so S-iridophores only change from loose to dense when there are no melanophores in the short range and simultaneously no xanthophores in the long range. In this case stripe integrity is lost in WT.

To summarise, all interactions are necessary for pattern formation in WT and single cell mutants, *nac* and *pfe*.

Biological redundancy

As part of this in-depth study, we have incorporated all of the interactions that we have identified from the literature. Consequently, there may be some in-built redundancy. However, we keep all interactions for the purposes of biological realism. In this section we explore the idea of biological redundancy by removing some interactions and observing the resultant simulated development.

First we consider movement. In real (and simulated) fish, melanocytes and xanthophores move 0.11 mm per day (***Takahashi and Kondo, 2008***) and 0.033 mm per day, respectively. Furthermore, in the short range their direction of movement is influenced by each other (***Yamanaka and Kondo, 2014***). Xanthophores chase melanocytes, which in turn, are repelled by xanthophores. In ***Figure 12C–D***, we turn off the movement of xanthophores and melanocytes in WT and *shd* mutants and simulate to stage J+. We observe that in both cases, the pattern is conserved. This is not surprising since the cells do not move very quickly. However, there is a slight difference in *shd* wherein the interstripe width is slightly smaller than in *shd* without changes (***Figure 12B***). We suggest that this is due to the loss of chase-run dynamics observed between melanocytes and xanthophores. Next we consider the differentiation rate of melanocytes. We model this as being dependent on the number of dense S-iridophores and the number of xanthophores currently on the domain. This is because we assume that dense S-iridophores and xanthophores positively influence the rate of

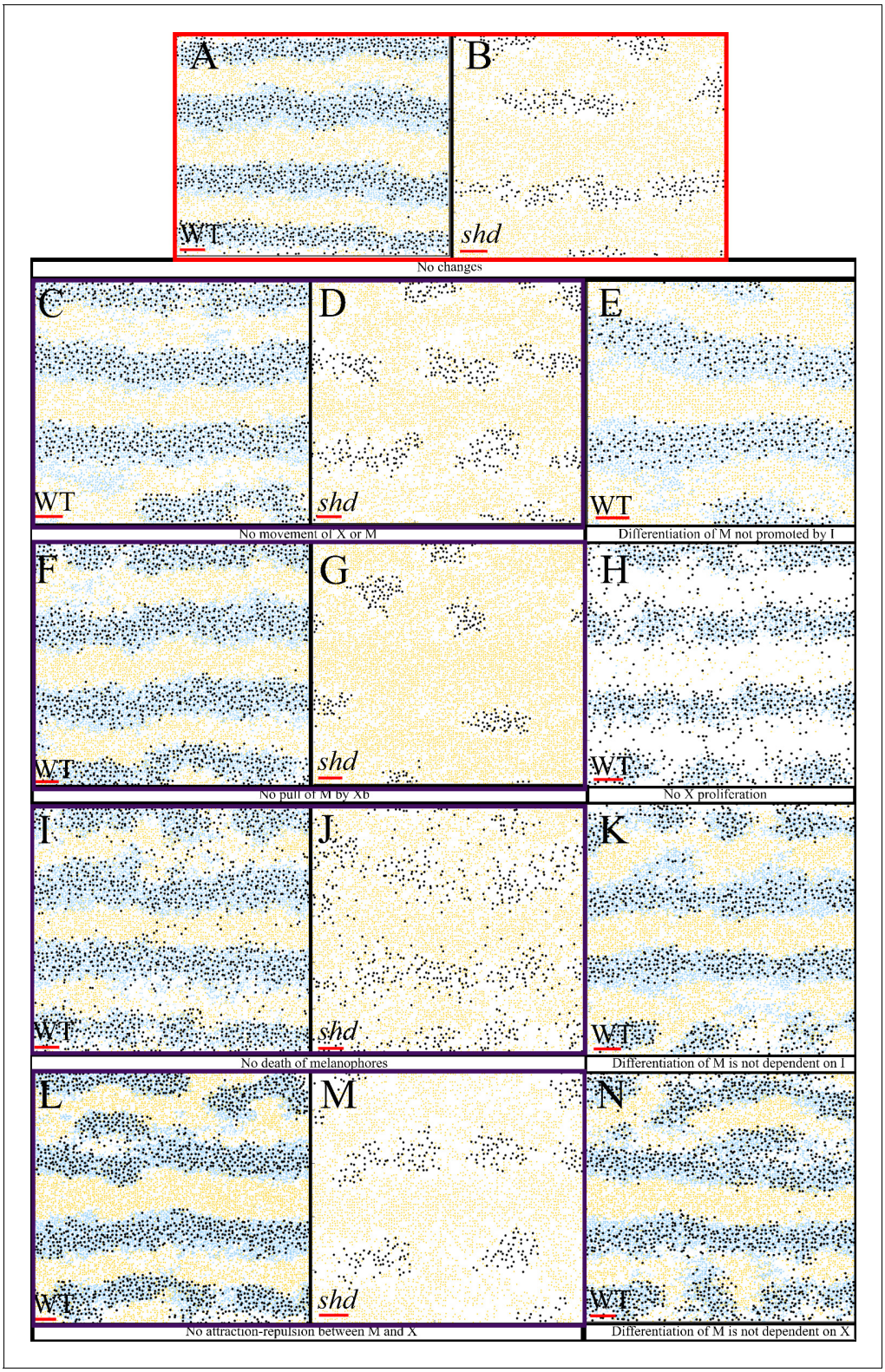


Figure 12. Representative simulation of the model with certain rules omitted. (**A–B**) WT and *shd* respectively with all rules included. (**C–D**) No movement of melanocytes or xanthophores for WT and *shd*, respectively. (**E**) Differentiation of melanocytes only promoted by dense S-iridophores for WT fish that is the rate of melanocytes does not depend on the number of S-iridophores. (**F–G**) No pull of melanocytes by xanthoblasts for WT and *shd* fish, respectively. (**H**) No proliferation of xanthophores for WT. (**I–J**) No melanocyte death for WT and *shd* fish. (**K**) Differentiation of melanocytes is not

Figure 12 continued on next page

Figure 12 continued

dependent on S-iridophores, (**L–M**) No death of melanocytes for WT and *shd* (**N**) Differentiation of melanocytes is dependent on dense S-iridophores being present in the long distance only.

melanocyte birth in the long range. In ***Figure 12E***, we change the differentiation rate so it is only dependent on the number of xanthophores currently on the domain and not the number of S-iridophores, effectively reducing the rate of differentiation of melanocytes in wild-type fish. The resultant pattern is still striped; however, it is less organised. In ***Figure 12F–G***, we remove the mechanism allowing long-range communication between xanthoblasts and melanocytes. In *shd* our model predicts that without the consolidation of spots by xanthoblasts, melanocyte spots become more widely spaced. In ***Figure 12H***, we remove xanthophore and xanthoblast proliferation. This limits the number of xanthophores to the number allocated at the start. Remarkably, our model predicts that stripe formation is largely preserved, however, interstripes are fainter due to the lack of xanthophores. Next we consider melanocyte death. In ***Figure 12I–J***, we remove melanocyte death from WT and *shd* simulations. Whilst stripe formation is maintained in WT (***Figure 12I***), interstripes are littered with melanocytes. In *shd* (***Figure 12J***), melanocyte spots are more difficult to discern as melanocytes can be observed in the xanthophore regions. We predict that dense S-iridophores and xanthophores in the long-range and/or loose S-iridophores and the lack of xanthophores in the short range promote melanocyte differentiation. In ***Figure 12K*** we change the criteria for melanocyte differentiation, so that it does not depend on S-iridophores (only xanthophores). As a result the stripe pattern looses integrity. A similar phenotypic change happens when we change the criteria for melanocyte differentiation, so that it does not depend on xanthophores (only S-iridophores, ***Figure 12N***). In ***Figure 12L,M***, we change melanocyte and xanthophore movement so that they are no longer influenced by each other (no chase-run dynamics). As a result in WT (***Figure 12L***), stripe integrity is lost. For *shd* (***Figure 12M***) simulations, however, qualitatively there is not much difference (similarly to the case when movement is completely removed (***Figure 12D***) suggesting that movement does not play a significant role in generating spots. Rather, it is the death of melanocytes in xanthophore rich areas and xanthoblast pulling which are important (***Figure 12G,J***).

In summary, whilst removal of these interactions largely do not change the type of pattern generated (e.g. *shd* still generates spots, wild-type fish still generate stripes), and thus could be considered as biologically redundant for pattern generation, they appear to have large impacts on the integrity of the patterns formed. We thus, suggest that the retention of these interactions in vivo act as a buffer to protect the integrity of stripe formation in spite of stochastic variations in stripe patterning.

Model predictions

A major benefit of developing a fine-grained, cell-level model is the ability to perform in silico experiments that can be directly related to real-life equivalents. This not only allows us to explore parts of the system that may otherwise not be testable experimentally, giving us valuable insight into the biological system. It also gives us the ability to analyse dynamics of different hypothetical mechanisms before devoting expensive resources to experimental tests that could confirm theoretical findings.

In the next few sections, we focus on the ability of the model to make biologically testable predictions, demonstrating firstly, in Section "An in silico investigation into important mechanisms for controlling pattern formation", how we can use our model to explore important facets of successful pattern formation such as growth, domain size and initial conditions. Then in Section "An in silico investigation into the function of the *leo* gene", we give an example of how we can use our model to generate testable hypotheses about the *leopard* mutant.

An in silico investigation into important mechanisms for controlling pattern formation

Initial S-iridophore interstripe orientation alone does not determine the orientation of stripes and interstripes

Previously it has been hypothesised that the horizontal orientation of the initial S-iridophore interstripe (emerging from the horizontal myoseptum) that drives the organisation of subsequent stripes and interstripes horizontally. One way to test this hypothesis is to initialise the interstripe so it is oriented vertically instead of horizontally. If the initial S-iridophore interstripe does orient stripes and interstripes then we would expect to see the same pattern development we observe in WT fish, but rotated 90 degrees. That is, we would expect to see vertical bars across the domain at the time corresponding to stage J+. The position of the dense S-iridophores (a horizontal interstripe along the horizontal myoseptum in WT fish) at the start of pattern metamorphosis, is dictated by the fish's anatomy and cannot be altered experimentally. However, we can simulate an altered iridophore initial distribution in silico by initialising the initial interstripe as a band of width three along the vertical axis instead of as a band of dense iridophores vertically (dorso-ventrally) down the centre of the domain of width three instead of as a band of width three, dense S-iridophores along the horizontal axis. We observe the subsequent pattern development at stages PR, SP and J+ in ***Figure 13A***. Interestingly, instead of observing vertical bars, at stage J+ we observe a labyrinthine pattern. This demonstrates that, whilst the initial S-iridophore interstripe plays a role in orientating the pattern, it is not the only part of the initial condition that is important. Further observation of the model output reveals that, for a while, the pattern is oriented in a vertical pattern similar to the initial iridophore interstripe, but as growth continues, this pattern becomes reoriented into a horizontal form. This clearly reveals the impact of growth on pattern formation.

The position of the initial S-iridophore interstripe is important for successful pattern formation

Another way to better understand the role of the initial S-iridophore interstripe is to alter its position in the dorso-ventral axis so that it appears more ventrally on the initial domain. We might naively predict, given that it has been hypothesised that dense S-iridophores only orientate the stripes and interstripes, that in this case the final pattern would be the same as in WT fish. We simulate this in silico by initialising the initial interstripe to be one quarter of the way up the domain instead of half way. A typical pattern evolution for this initial condition is displayed in ***Figure 13B***. We observe that subsequent stripes and interstripes still appear sequentially, either side of the initial interstripe, suggesting that the S-iridophores do play a role in the positioning of new stripes and interstripes. However, we do not observe usual WT patterning. In particular, stripes and interstripes exhibit more breaks compared to WT simulations. Moreover, developing stripes and interstripes become sequentially thinner as a result of the impact of domain growth. Once again, growth is the key factor: growth is centred at the middle of the domain and so when the initial stripe is not similarly centred, growth disrupts pattern formation in our model.

Initial domain size contributes to the number of stripes and interstripes

In order to test the role of domain size in pattern development, we initialise the domain so that it is three times as tall as in WT simulations. That is, we initialise the domain to be 2 mm $\times$ 3 mm instead of 2 mm $\times$ 1 mm. All other parameters remain unchanged, including the rate of growth in the horizontal and vertical axis. We present a typical example of subsequent pattern development in ***Figure 13C***. At stage PR, pattern development is similar to the pattern seen in WT fish at the same stage. However, at stage SP, we observe more stripes than are observed even by stage J+ in WT fish. By stage J+, instead of three interstripes and four stripes as seen in WT, we observe six stripes and seven interstripes. This suggests that the initial domain height influences the number of stripes and interstripes that develop, provided that growth is uniform and centred.

Stripe insertion can occur on an initially striped domain

Kondo and Asai observed that, as the size of the marine angelfish *Pomocanthus* doubled, new stripes along the skin would develop between the old ones (***Kondo and Asal, 1995***). This phenomenon has not been observed in zebrafish, where new stripes and interstripes appear consecutively at the dorsal and ventral periphery. We hypothesise that this is likely related to either pattern

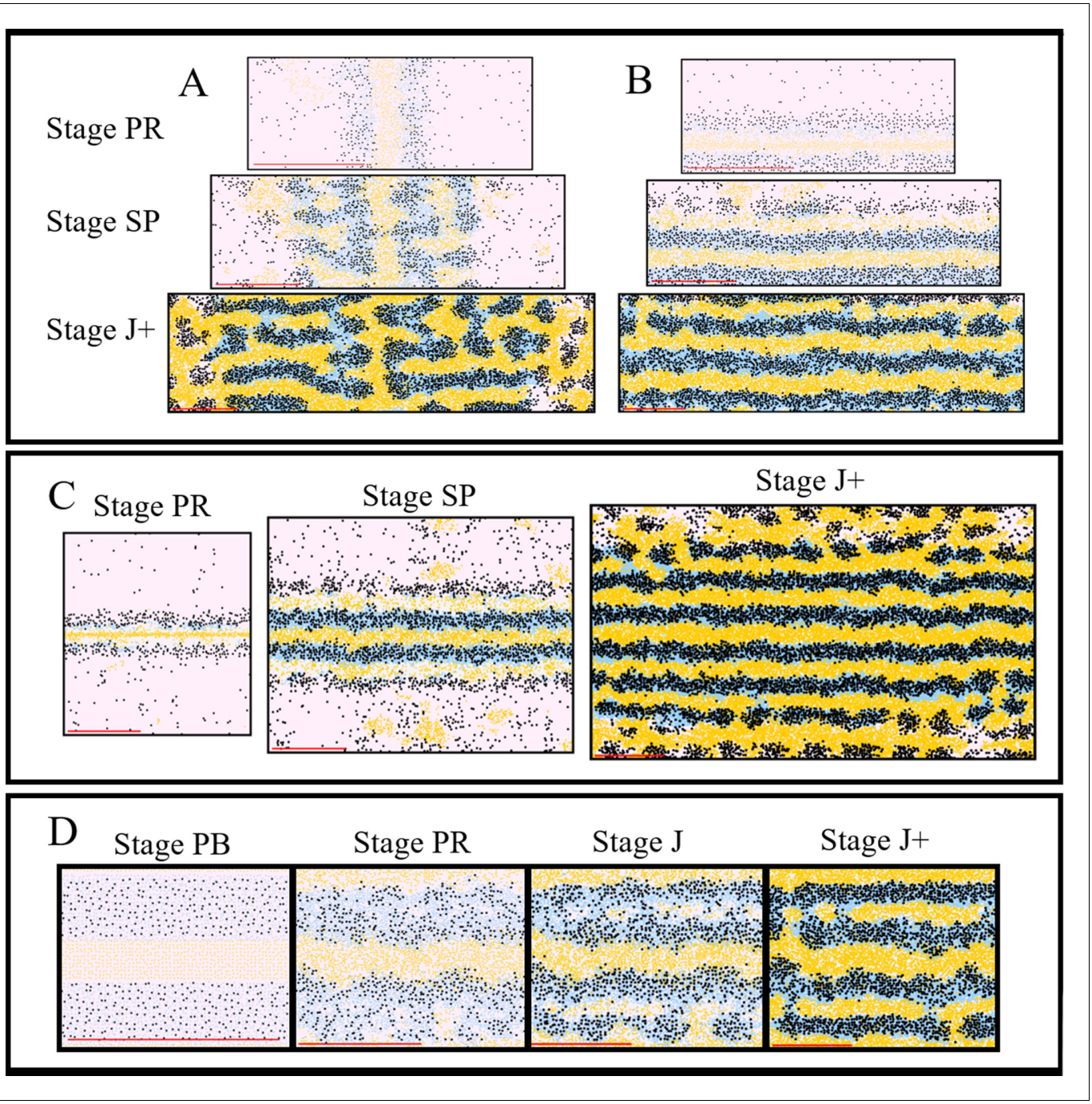


Figure 13. In silico investigation into important mechanisms for controlling pattern formation. The simulated domains at stages PR, SP and J+ wherein the following are changed (**A**) the orientation of the initial S-iridophore interstripe, (**B**) the initial position of the S-iridophore interstripe, (**C**) the initial domain size (**D**) the initial domain so that it is populated with adult-width stripes and interstripes.

maintenance mechanisms or the spatial localisation of growth. Here, we experiment with the model to see whether stripe insertion can occur when the domain is populated at stage PB with adult-width stripes and interstripes. The results of an example realisation with these initial conditions are given in ***Figure 13(D)***. We observe that, in this case, new interstripes do appear between pre-established stripes. This is because growth (which is centred in the middle of the domain) creates space within the middle of the already developed stripes and interstripes.

S-iridophores are more important to the generation of melanocytes than xanthophores

We also used the model to make some more subtle predictions. For example, in the case of melanocyte differentiation, which we model as being promoted in the long range by both xanthophores (from observation of ablation experiments; ***Kondo and Asal, 1995***) and S-iridophores (from observations of *pfe*), there were no known parameter values for their relative strengths. We found using our model that by making the strength of S-iridophore promotion of melanocyte differentiation to be much greater than that of xanthophores, qualitatively and quantitatively the model simulated for WT, *pfe* and *shd* was greatly improved (see ***Figure 14A–B***).

Horizontal growth bias during development generates more tortuous stripes in WT fish

Interestingly, we also observed in our simulations that increased height-to-length ratio is correlated with stripes becoming more tortuous ($R = -0.617$, $p<0.01$, ***Figure 14C***). This phenomenon is not

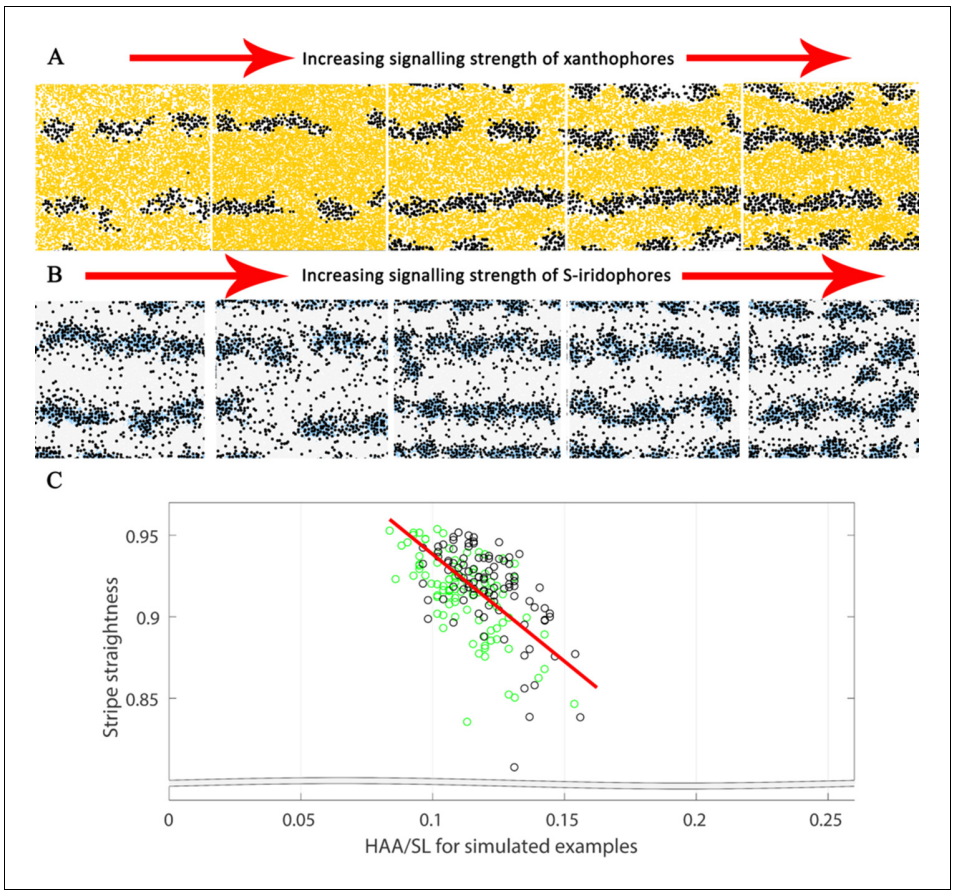


Figure 14. Substantial alterations in starting conditions generate major alterations in pattern formation. The effects of different signalling strengths in *pfe* and *shd*. Representative images of J+ simulations of (**A**) *shd* and (**B**) *pfe* where the signalling strength of melanocyte differentiation in the long range is increased (from left to right). Increasing the signalling strength of the respective cell types decreases the width of the interstripe X0 and increases the number of pseudo-stripes from 2 to 4. For our simulations, we choose the weakest shown signalling strength for xanthophores, corresponding to the furthest left of (**A**) and the strongest shown signalling strength for S-iridophores, corresponding to the furthest right of (**B**). (**C**) Stripe straightness of our simulations correlates with the simulated HAA/SL ratio. Each circle is one simulation, green circles indicate simulations up to stage J, black circles indicate simulations of up to stage J+. There are one hundred simulations in total for each stage. The red line is the line of best fit.

something we can see as being consistent with real fish and thus suggests that some interactions may be missing regarding the maintenance of stripe and interstripe formation.

An in silico investigation into the function of the *leo* gene

The model provides testable hypotheses for cryptic functions of the *leo* gene

The gene *leo* encodes Connexin39.4 (Cx39.4) (**Watanabe et al., 2006**; **Maderspacher and Nüsslein-Volhard, 2003**; **Irion et al., 2014**). As a result, *leo* mutants display a leopard-like spotted pattern across the flank of the fish (**Figure 15A–A'**), instead of the usual striped pattern (**Figure 1A**). In this section, we aim to hypothesise key aspects of the *leo* mutations using our model alongside observations of relevant mutants.

Pattern formation is also altered in the double mutants *leo;shd*, *leo;nac* and *leo;pfe* when compared with *shd*, *nac* and *pfe* (**Irion et al., 2014**). For example, the flank of double mutant *leo;nac* is covered by xanthophores and dense S-iridophores (**Figure 15B''**). This is in contrast to *nac* which also contains large patches of loose S-iridophores (**Figure 15B–B'**). Adult *leo;pfe* fish exhibit randomly distributed melanocytes instead of spots (**Figure 15C–C''**). Finally, adult *leo;shd* exhibit an absence of melanocytes on the flank of the fish, instead, the flank of the fish is entirely covered with xanthophores. This is contrast to the melanocyte spots normally observed on *shd* (**Figure 15D–D''**).

Connexins are involved in cell–cell communication and signalling. Since, Cx39.4 is required for normal function in melanocytes and xanthophores but not in S-iridophores (**Watanabe et al., 2006**; **Maderspacher and Nüsslein-Volhard, 2003**; **Irion et al., 2014**), this suggests that in *leo*, cell–cell communication between melanocyte and xanthophores may be disrupted. Moreover, from observation of the double mutants, it seems that *leo* presumably generate heteromeric gap junctions among and between melanophores and xanthophores, controlling S-iridophore shape transitions (**Irion et al., 2014**).

In order to investigate the influence of the *leo* gene, we first consider the individual cell–cell interactions that can be deduced from the literature and ask if these cell–cell interactions are sufficient for generating the pattern. So far, there has been one experimental study observing the individual behaviour of *leo* cells. **Kondo and Watanabe, 2015** studied the movement in-vitro of *leo* melanocyte and xanthophore cells. They demonstrated that the *leo* melanocyte repulsive response to xanthophores was hardly observed in comparison to the marked repulsion in WT fish. This suggests that melanocyte repulsion from xanthophores is inhibited in *leo* (**Kondo and Watanabe, 2015**). We will refer to this as hypothesis one for the effects of the mutant *leo*.

- Hypothesis 1:Melanocytes are not repelled by xanthophores.

We can simulate pattern development in this case by turning off melanocyte repulsion by xanthophores in our model. This is shown in **Figure 15E** in the column numbered 1. We observe that at J+ the pattern consists of thicker interstripes than in WT fish, but not spots. Simulating the *shd* phenotype (lack of S-iridophores) with hypothesis 1, also does not generate the pattern expected in *leo;shd*. Hypothesis one is also insufficient to explain the phenotype of *leo;nac* or *leo;pfe*. This is because there are either no xanthophores or no melanocytes respectively in these mutants and therefore no melanocyte-xanthophore interactions. From these observations, we conclude that hypothesis one alone is insufficient for *leo* pattern formation.

To deduce the other cell-cell interactions that a mutation in *leo* could affect, we look at the patterns of adult *leo;nac* (**Figure 15B''**) and *leo;shd* (**Figure 15D''**). Of the three double mutants, these are the easiest from which to deduce single cell-cell interaction changes that could generate the double mutant patterns.

The *leo;nac* mutant displays an absence of loose S-iridophores compared to *nac*, suggesting that signalling of xanthophores which promote S-iridophore transition from dense to loose is inhibited in *leo*. The *leo;shd* mutant displays an absence of melanocytes in the adult pattern compared to *shd*, suggesting that long-range survival signal sent by xanthophores to melanocytes is inhibited in *leo*. We propose two further potential hypotheses for the effects of the mutant *leo*.

- Hypothesis 2: Xanthophores do not promote the survival of melanocytes in the long range.
- Hypothesis 3: Xanthophores do not promote the change of S-iridophores from dense to loose in the long range.

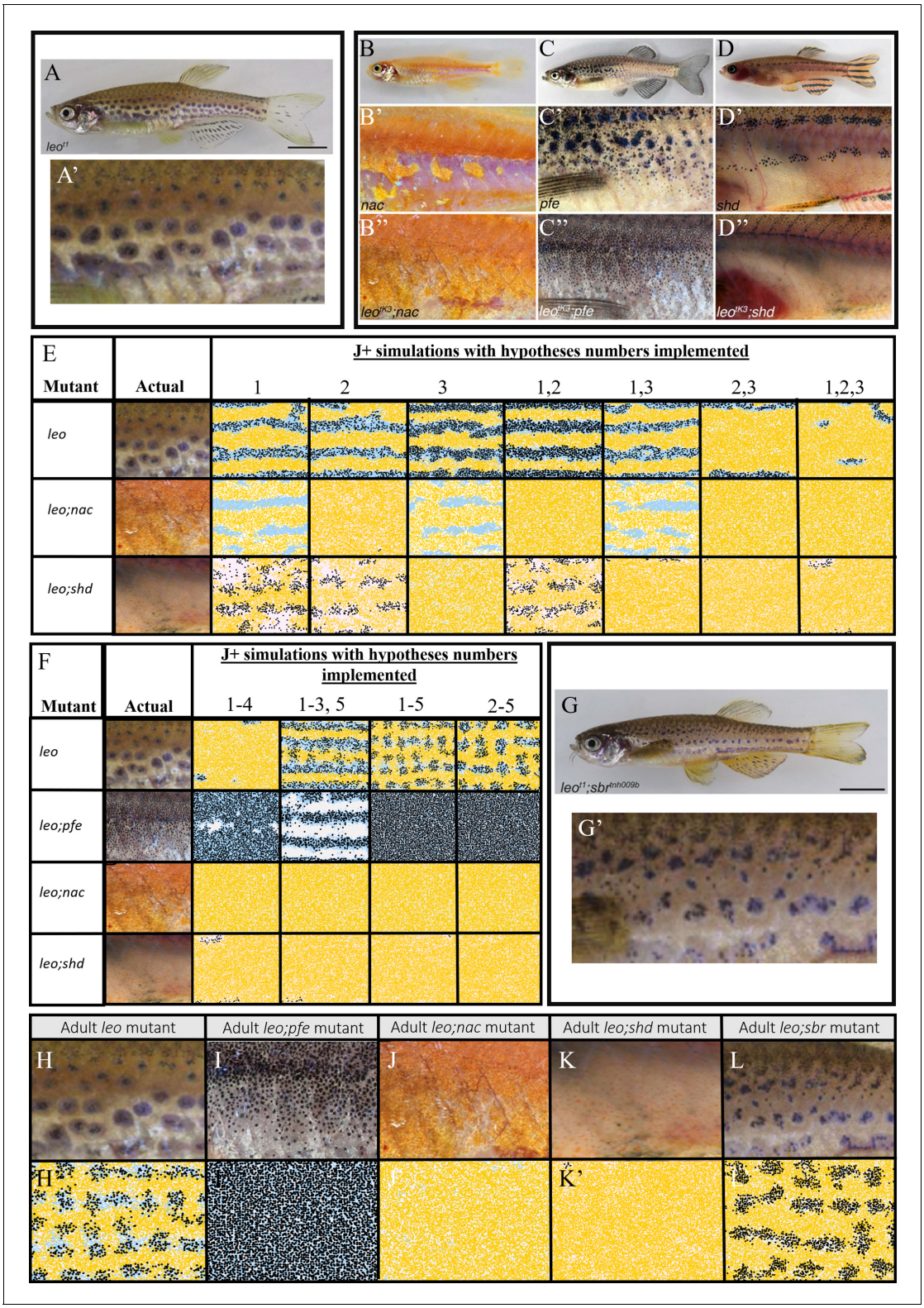


Figure 15. Using the model to generate predictions regarding the underlying effects of the *leo* mutant. (**A–D**) Adult *leo*, *nac*, *pfe* and *shd* respectively. (**A'–D'**) The flanks of *leo*, *nac*, *pfe* and *shd* respectively. (**B''–D''**) The flanks of *leo;nac*, *leo;pfe* and *leo;shd* respectively. (**E**) Example simulations where combinations of hypotheses 1–3 in the text are implemented. (**F**) Example simulations where combinations of hypotheses 1–5 in the text are implemented. Note that simulations of 1–5 are similar to simulations for 2–5, indicating that hypotheses one is not necessary to predict the *leo*

Figure 15 continued on next page

Figure 15 continued

phenotype. (**G**) Adult *leo;sbr* (**G'**) Flank of *leo;sbr*. (**H–L**) Adult mutants contrasting with (**H'–L'**) simulations using hypotheses 1–5 from the main text. (**H–L**) Flank of *leo*, *leo;nac*, *leo;pfe*, *leo;shd* and *leo;sbr* respectively. (**H'–L'**) Simulated images of *leo*, *leo;pfe*, *leo;nac*, *leo;shd* and *leo;sbr* respectively at stage J+. (**A**), (**A'**), (**G**), (**G'**), (**H**) and (**L**) from ***Fadeev et al., 2015***, (**B–D''**), (**I–K**) from ***Irion et al., 2014*** and are all licensed under CC-BY 4.0 (http://creativecommons.org/licenses/by/4.0).

In ***Figure 15E***, we provide a table of results for all combinations of hypotheses 1–3. We comment that hypotheses 1–3 cannot be all-encompassing, as so far, none of our hypotheses effect melanocyte-iridophore interactions and thus cannot generate the phenotype *leo;pfe*. Meanwhile, we focus on being able to generate *leo;nac* and *leo;shd*.

In ***Figure 15E***, we demonstrate, that none of the hypotheses alone can generate the *leo* pattern, nor can they generate more than one of the double mutant types. When hypotheses 1 and 2 are combined, striping is disrupted, there are more breaks than in WT and interstripes are wider than WT. When 1 and 3 are combined, stripes are the same as in WT, presumably due to compensation of other cell-cell interactions, however, the simulated *shd* pattern has less aggregation of melanocytes than typically observed in *shd*. When hypotheses 2 and 3 are combined we observe the expansion of the initial interstripe to the very edges of the domain dorso-laterally, where there are some melanocytes, unlike *leo*. However, in this case, our simulations of *leo;nac* and *leo;shd* match the real phenotype. We demonstrate that when all hypotheses are implemented we generate a unstable pattern of spots which, eventually disappear over time, leaving a domain consisting of dense S-iridophores and xanthophores. Also, we can replicate *leo;nac* and *leo;shd*.

Finally, we attempt to elucidate the melanocyte-iridophore cell-cell interactions that are affected in the *leo* mutant by considering the phenotype of adult *leo;pfe* (***Figure 15C'–C''***). The *leo;pfe* mutant displays a random distribution of melanocytes below a field of S-iridophores, suggesting that *leo* affects directed differentiation of melanocytes by dense S-iridophores. This leads us to two extra hypotheses for the effects of the *leo* gene:

- Hypothesis 4: Melanocytes lose death signals from local dense S-iridophores and, as a result, can differentiate in dense S-iridophore zones.
- Hypothesis 5: Melanocytes lose directed signalling from S-iridophores and hence, in the absence of xanthophores differentiate randomly.

In ***Figure 15F***, we provide a table of representative simulated patterns for all combinations of hypotheses 1–3 with hypotheses 4 and 5. Any of the combinations considered can generate the *leo;shd* and *leo;nac* mutant phenotypes. Hypotheses 1–4 and 1–3,5 cannot generate the *leo* spots. However, if we combine all five hypotheses then we can replicate the phenotypes of all considered mutants; *leo*, *leo;pfe*, *leo;nac* and *leo;shd*. These results suggest that hypotheses 1–5 are sufficient for explaining *leo*.

Necessity of the pre-existing hypothesis about *leo*

Next, we evaluate the necessity of hypothesis (1) Hypothesis 1 is the only assumption that comes directly from the literature (***Yamanaka and Kondo, 2014***). We show, using our model, in ***Figure 15F*** that there is no notable difference between the patterns generated with and without hypothesis 1 (i.e. hypotheses 1–5 versus hypotheses 2–5) other than that the spots may be slightly more irregular in when hypothesis 1 is included. This suggests that the cell–cell interaction mediating repulsion of melanocytes by xanthophores in *leo*, may not be necessary to generate the characteristic spots.

Hypotheses 1–5 can replicate the patterns of *leo*, *leo;nac*, *leo;pfe* and *leo;shd*

Figure 15H—K display the flanks of *leo*, *leo;pfe*, *leo;nac* and *leo;shd* respectively. ***Figure 15H'–K'*** display the simulated patterns at stage J+ for in silico mutants *leo*, *leo;pfe*, *leo;nac* and *leo;shd* respectively where our assumptions for the function of *leo* are based on hypotheses 1–5 and, in the case of double mutants *leo;pfe*, *leo;nac* and *leo;shd*, our assumptions for *pfe*, *nac* and *shd* are as previously described in Section "Simulation of 'missing' cell mutants".

We observe, for our simulations of *leo* (***Figure 15H'***), that melanocyte spots are associated with loose S-iridophores in a sea of dense S-iridophores and xanthophores just as in real *leo* fish (***Figure 15H***). For our simulations of *leo;nac* (***Figure 15I'***) we observe a domain that is fully populated with xanthophores and dense S-iridophores as expected (***Figure 15I***). Our simulations of *leo;pfe* (***Figure 15J'***) is fully populated with S-iridophores and randomly distributed melanocytes as expected (***Figure 15J***). Finally, our in silico representation of *leo;shd* (***Figure 15K'***) is fully populated with xanthophores only, as seen in ***Figure 15K***.

The *leo;sbr* cross mutant phenotype is an emergent property of the model

Previously in Section "Simulation of other mutants" we considered the *sbr* mutant. The *sbr* gene encodes Tight Junction Protein 1a (Tjp-1a), which is expressed cell autonomously in dense S-iridophores and causes the shape transition from dense to loose to be delayed (***Fadeev et al., 2015***). We showed in ***Figure 7K'–N'*** that our model could replicate the *sbr* pattern development by altering the rate at which S-iridophores attempt to change from loose to dense to be slower than in WT.

Figure 15G—G' shows a typical example of the *leo;sbr* phenotype. The adult *leo;sbr* displays spots which, compared to *leo* (***Figure 15A–A'***), are more elongated. ***Figure 15L*** displays the flank of an adult *leo;sbr*. In ***Figure 15L'***, we simulate a mutant that satisfies hypotheses 1–5 as well as our assumptions about *sbr* to generate an in silico leo;sbr. Consistent with real *leo;sbr* mutants, our simulation of *leo;sbr* has melanocyte spots associated with loose S-iridophores that are surrounded by xanthophores and dense S-iridophores. Comparison with our simulation of *leo* (***Figure 15H'***) demonstrates that our simulation of *leo;sbr* also displays more elongated spots than those of our *leo* simulations.

Robustness of the *leo* assumptions in generating spots

As a test of robustness, we perform a rigorous robustness analysis by carrying out one hundred repeats of the mutant simulations with perturbed parameter values chosen uniformly at random from the range 0.75–1.25 of their described value as in Section "Simulation of WT pattern". Ten of these randomly sampled repeats are given in Figure ***Appendix 4—figure 1*** for *leo*. We observe that for all one hundred repeats that small perturbations to the rates still generate consistent spots, demonstrating the robustness of the model.

These results of this section demonstrate the remarkable ability of our model to generate the *leo* single and double mutant phenotypes under a set of specific proposed changes to the model rules; these proposals can be used to guide experimental exploration of the effects of the *leo* gene.

Discussion

As a result primarily of beautiful experimental work, zebrafish pigment pattern formation has become the best characterised pigment patterning mechanism (***Frohnhöfer et al., 2013***; ***Kelsh et al., 2009***; ***Singh and Nüsslein-Volhard, 2015***; ***Watanabe and Kondo, 2015***; ***Patterson and Parichy, 2019***). Zebrafish pigment pattern formation requires three pigment producing cell types: melanocytes, xanthophores and S-iridophores to generate the WT stripes (***Frohnhöfer et al., 2013***). In an attempt to decipher the mechanisms underlying pattern formation, most previous mathematical models have largely focussed on xanthophores and melanocytes and neglected S-iridophores (***Nakamasu et al., 2009***; ***Bullara and De Decker, 2015***; ***Volkening and Sandstede, 2015***; ***Painter et al., 2015***; ***Bloomfield et al., 2011***). However, recent studies have shown that S-iridophores are a crucial component, with S-iridophore transitions between dense and loose driving pattern formation. First, the early provision of dense S-iridophores through the horizontal myoseptum orients stripes (***Frohnhöfer et al., 2013***; ***Kelsh et al., 2009***; ***Singh and Nüsslein-Volhard, 2015***; ***Watanabe and Kondo, 2015***; ***Patterson and Parichy, 2019***). Moreover, S-iridophores sequentially pre-determine stripes and interstripes via respective shape changes in these regions (***Frohnhöfer et al., 2013***; ***Fadeev et al., 2015***; ***Krauss et al., 2014***).

Here, we have taken a bottom up approach to modelling zebrafish pattern formation with the aim of testing whether the experimentally defined set of biological rules for zebrafish pigment pattern formation might be sufficient to explain both the WT and the diversity of mutant pigment patterns. We used an individual based modelling approach incorporating all five cell-types deemed important for pattern formation in zebrafish. We formalised all respective cell-cell interactions

mathematically, with interaction strengths, parametrised, where possible, by the biological literature (see ***Supplementary file 5***). For the less well-studied S-iridophore transitions, we analysed key mutant phenotypes to infer biologically realistic rules for these interactions, aiming to generate assumptions that were the simplest for pattern formation changes seen, but no simpler. We proved our models ability to simulate the distinctive pattern features during developmental stages PB through J+ of each of WT and six mutant patterns that had been used to determine the biological rules. We showed that in each case, our model simulations matched qualitatively the pattern development in real fish at the various developmental stages considered. This is consistent with the proposal that our modelling assumptions were sufficient for pattern development in these cases. As a more rigorous test of the model, we then investigated its ability to successfully simulate the distinctive patterns of three further mutants with defective S-iridophore properties, including two mutants that had not been modelled mathematically before. We showed that in each case, our model also correctly replicated patterns that were qualitatively similar to the corresponding mutant fish at various developmental milestones.

We assessed multiple quantitative features of our simulations against measured data from published studies, focusing on spatial distributions of cell-types, stripe width and melanocyte numbers. We found that in each case our simulations were highly reproducible, and quantitatively matched the biological observations. We conclude that our mathematical modelling approach, built upon the biological literature, provides substantial validation of the sufficiency of that set of biological rules in explaining pattern formation in zebrafish development for WT and many other mutant fish. Furthermore our modelling provides support for the plausibility of the deduced rules for S-iridophore packing transitions during pattern development.

Finally, we demonstrated the capability of our model to give valuable insight into the patterning mechanism and to make testable predictions about the biology.

This paper represents the first demonstration, to the best of our knowledge, of a model being used explicitly to test the impact of the *sbr* mutation. The *sbr* gene encodes Tjp1a, a key tight junction protein, and is expressed at much higher levels in S-iridophores in a dense configuration than those of the loose form (***Fadeev et al., 2015***). Furthermore, double mutant and chimaeric studies show that *sbr* acts cell-autonomously within the S-iridophores to control adult pigment pattern formation (***Fadeev et al., 2015***). These authors also show that in *sbr* mutants the transition from dense to loose S-iridophores is delayed, suggesting that this transition somehow depends upon Tjp1a in dense S-iridophores (***Fadeev et al., 2015***). Here, we test the patterning impact of this interpretation, by incorporating delayed S-iridophore state transition into our model, and show that this does indeed result in pattern changes consistent with the *sbr* phenotype. This provides theoretical support for Fadeev and colleagues deductions and deepens the interest in understanding the mechanistic basis for this role for Tjp1a.

Our modelling results demonstrate the applicability of complex models to test hypotheses that are difficult to test experimentally. Previously, it has been hypothesised that S-iridophores contribute to pattern formation by orienting the stripe. In Section "An *in silico* investigation into important mechanisms for controlling pattern formation", we demonstrate that this is true. We show that simply reorientating the interstripe to a vertical position is not enough to produce vertical bars as our simulations exhibit a labyrinthine pattern instead of vertical bars. Careful observations of our simulations indicate that in addition to the initial condition, growth is important for determining the final pattern. We show that moving the initial position of the interstripe away from the centre of the flank, where growth is centred, also disrupts the patterning. We are also able to show that by enlarging the initial domain so that it is the same width but taller in height we show that we can generate more stripes and interstripes than that are usually observed at stage J+. Finally, by initialising a domain that is already populated with cells in a stripe position that we can replicate a different stripe formation mechanism seen in fish *Pomocanthus*. Whilst in zebrafish stripe formation is sequential, starting from an initial interstripe and developing bidirectionally from the middle, *Pomocanthus* develops stripes in-between other stripes. We show that if a striped pattern is fully formed when the fish is still growing (and growth is centred) that this forces new stripes to occur between the old ones, instead of at the periphery.

Our model is the first, to the best of our knowledge, to suggest that S-iridophores play a more important role in melanocyte differentiation than xanthophores. We found that by implementing a stronger signaling capacity of S-iridophores than xanthophores for long range melanocyte

differentiation then we obtained a better qualitative and quantitative match for *shd* and *pfe* mutant patterns. In particular, a stronger signalling success rate for melanocyte differentiation from S-iridophores in the long range appeared to align subsequent stripes in *pfe* and WT, resulting in consistency of pseudo-stripe and stripe width respectively. The comparatively reduced signalling success rate for melanocyte differentiation from xanthophores in the long range reduced the number of pseudo-stripes in *shd* from four, to two at stage J+ which is more consistent with real data. Real *shd* fish typically exhibit fewer stripes (approx two at stage J+) than the four of WT (**Frohnhöfer et al., 2013**; **Figure 4A–B**). We also found that this factor was important to produce melanocyte numbers that quantitatively matched data by Frohnhöfer et al when comparing between mutant and WT (**Figure 10A**) as well as a better stripe width match (**Figure 10B**). Thus, our study further reinforces findings that S-iridophores play an important role in determining stripe and interstripe width (without S-iridophores, X0 interstripe width in *shd* is increased) and not just the widely reported role of stripe alignment.

We have further built upon previous mathematical modelling work, by using our model to make predictions about the functions of the *leo* gene. The gene *leo* encodes Connexin 39.4, which is required in melanocytes and xanthophores but not S-iridophores (**Irion et al., 2014**). Connexins play an important role in cell–cell signalling and communication so it has been postulated that the *leo* gene is involved in the signalling between melanocytes and xanthophores as well as signalling cues to S-iridophores regarding shape-transitions (**Irion et al., 2014**). Previous to our investigation, there had been one study of the individual cell-cell interactions in *leo*. This study by demonstrated that unlike WT melanocytes, *leo* melanocytes are not repelled by xanthophores in the short range (**Yamanaka and Kondo, 2014**). Using our model, we first demonstrated that this cell-cell interaction alone was not enough to reproduce the *leo* mutant pattern. Then, in a systematic approach we deduced four hypotheses about cell–cell interactions that might be affected by *leo*, which, upon implementing in our model, successfully replicated the patterns observed in *leo, leo;pfe, leo;nac, leo;shd* and *leo;sbr*. This work provides testable hypotheses about the effect of *leo* which can now guide future experimental work.

In contrast to most previous mathematical studies of pattern formation, the rules we propose for zebrafish pigment patterns are complex and extensive. For example, successful S-iridophore shape transitions in our model require information from both melanocytes and xanthophores. Many other studies have condensed zebrafish pattern formation to a few simple rules that can often be described by a series of partial differential equations (**Kondo, 2017**; **Painter et al., 2015**; **Nakamasu et al., 2009**), in particular Turing reaction-diffusion-type models posit that combinations of short and long range dynamics between melanocytes and xanthophores generate stripe patterns. Indeed, a lot of the excitement around such models is the ease with which small parameter value changes can sometimes result in diverse patterns, many readily recognisable from nature (**Watanabe and Kondo, 2015**; **Metz et al., 2011**; **Maini et al., 2012**). A major difference between our model and Turing reaction-diffusion models is that small parameter changes in our model do not typically generate qualitatively different patterns, whereas Turing reaction-diffusion models can show substantial pattern changes in response to small alterations to parameters near to bifurcation points (**Watanabe and Kondo, 2015**; **Maini et al., 2012**). We suggest that the added complexity of the real system has evolved to make the patterning process robust, with partially redundant mechanisms insulating against the impact of stochastic variation during pattern formation.

Our modelling approach is analogous to that adopted in another recent study of zebrafish pattern formation, one which independently attempts to understand S-iridophore contributions to the process (**Volkening and Sandstede, 2018**). In a similar fashion, Volkening et al generated an off-lattice model incorporating S-iridophores for which the rules were based upon the experimental literature. Upon implementing these rules they attempted to test the sufficiency of the known biology in describing the patterning process. Importantly, Volkening et al's modelling approach differs from that adopted here in several ways. First, we use an on-lattice model, whilst Volkening et al use an off-lattice model. Using an on-lattice model allowed us to incorporate volume exclusion by other cells directly. Secondly, we used a continuous-time model, whereas Volkening et al update their model at simulated 24 hr intervals, with all rules implemented simultaneously. By using a continuous-time method, we are able to capture the stochasticity involved in rates of reactions and the ordering of events over time. Finally, we incorporate a hypothesis that S-iridophores contribute more strongly to promoting melanocyte differentiation than xanthophore differentiation, leading us to a better

qualitative approximation of *shd* mutants and a better quantitative estimate of melanocyte numbers in *shd*, *pfe* and *rse* than shown before.

Importantly, the model of Volkening et al. also proves highly capable in generating simulations that accurately mimic WT and various mutant pigment patterns. This observation further strengthens support for the validity of the proposed S-iridophore rules we each postulate. More broadly, we consider that our independent mathematical approaches are mutually reinforcing in reaching the conclusion that the deduced biological rules may be largely sufficient to explain pigment pattern formation in zebrafish.

Further testing of our model should focus on investigating later stages of pigment pattern development and maintenance. To date, our focus has been on the crucial dynamic development between PB and J+ stages when the pattern is evolving; we have not considered pattern maintenance between J+ and adulthood. Work by **Frohnhöfer et al., 2013** has demonstrated that in *pfe* mutants, which up to stage J+ have similar numbers of melanocytes to WTs (approximately 90% of WT numbers), this drops to approximately 50% of WT by adulthood. Whilst our model correctly predicts melanocyte numbers in *pfe* compared to WT prior to J+, preliminary simulations up to adulthood suggest this sharp decrease in melanocyte numbers shown in real fish are currently not predicted by our model. This clearly indicates that new biological mechanisms concerning maintenance, likely involving the late differentiating L-iridophores, need to be analysed and incorporated into a model that extends to these later developmental stages.

From an analytical perspective, an advantage of our on-lattice model is its amenability to the derivation of a continuum model, although we note that continuum approximations to off-lattice individual-based models can also be derived. Our model therefore opens up the opportunity for future exploratory work using a continuum model for mutants *pfe* and *nac* in order to explore whether pattern formation in these cases individually can be described as Turing patterns and to determine parameter ranges for successful pattern formation.

It will also be interesting to investigate the role of growth in pattern formation and maintenance. We observed in our simulations a lower average stripe straightness (0.92) than to that of real WT fish (0.98). We further observed in our simulations that increased height-to-length ratio is correlated with stripes becoming more tortuous (**Figure 14C**). Stripes in real fish seem, qualitatively, to not show this effect, and we suspect that our model can be further refined here. A search for mechanisms that might increase stripe straightness will be valuable here, and we note that exploration of our model should allow candidate mechanisms to be identified in silico. Further investigation of this feature in our model once extended to adult pattern maintenance may also be important. For example, casual observation suggests that fish that show a particular height to length bias do not show notably tortuous stripes. Therefore, future work will be to understand what preserves the pattern in these cases. We predict that this may be related to the position of growth. Thus, future work will be to fully evaluate the effects of growth on the final pattern formation.

Another significant aspect which deserves attention concerns dorso-ventral pattern differences. In fish this is characterised by having more pigmented cells at the dorsal region than the ventral region. This is certainly true in adult *nac* (**Figure 4C**), which has disproportionately more pigmented xanthophores in the dorsal than the ventral region). We have recently noted the subtle impact of dorso-ventral countershading on the WT zebrafish pigment pattern, including in the stripes themselves, and have identified Agouti as a key regulator of this process (**Cal et al., 2019**). Furthermore, there are clear dorso-ventral asymmetries in some of the adult mutant patterns: e.g, *nac* mutants exhibit a strong X0 interstripe and a weak ventral interstripe X1D, and completely lack dorsal interstripes. Our model will allow us to explore possible drivers of this asymmetry, which we hypothesise will include Agouti signalling and also differential domain growth.

In conclusion, our on-lattice model, implementing the current biological understanding of adult zebrafish pigment pattern formation, strongly supports the validity of these experimental interpretations, motivating the detailed investigation of their molecular bases. Our model also highlights areas where knowledge is currently incomplete and, importantly, has allowed in silico investigations to identify plausible mechanisms that require experimental testing.

Additional information

Funding

Funder	Grant reference number	Author
Biotechnology and Biological Sciences Research Council	SWBio DTP	Jennifer P Owen
Biotechnology and Biological Sciences Research Council	BB/ L00769X/1 (RNK)	Robert N Kelsh

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Jennifer P Owen, Conceptualization, Data curation, Formal analysis, Investigation, Methodology; Robert N Kelsh, Resources, Supervision, Investigation, Methodology; Christian A Yates, Supervision, Investigation, Methodology, Project administration

Author ORCIDs

Jennifer P Owen https://orcid.org/0000-0001-8440-6822
Christian A Yates https://orcid.org/0000-0003-0461-7297

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.52998.sa1
Author response https://doi.org/10.7554/eLife.52998.sa2

Additional files

Supplementary files

- Supplementary file 1. Summary of all notation used in Appendix 1.
- Supplementary file 2. Summary of all notation used in Appendix 1 regarding short and long range interactions.
- Supplementary file 3. SL measurements and HAA measurements and WT pattern description by stage are given as by ***Parichy et al., 2009***. Corresponding dpf for the stages are approximated from the images of ***Frohnhöfer et al., 2013***.
- Supplementary file 4. All continuous time events and their corresponding propensities for event attempts at time t, denoted $\alpha_i(t)$ for event i in minutes.
- Supplementary file 5. Full description of all the cell–cell interaction given in main text. Melanocytes, xanthophores, loose iridophores, dense iridophores and xanthoblasts are denoted M, X, I^l, I^d, X^b respectively. C1, C2 stands for cell 1, cell 2, where $C1 = X$, X^b is the target cell and cell two is the signalling cell with corresponding signal range (R) that is either short (S), up to 0.04 mm, or long (L), 0.12 mm. This signal generates action (A) of cell type C1 of: movement (M), differentiation (D), proliferation (P) or survival (S) of type (T). For action (D) or (S), type (T) is denoted '+' if the resultant action is promotion of action A and '-' if the resultant action is inhibition of action (A). For action (M), type (T) is denoted '+' if the resultant action is attraction towards cell type C1 and '-' if the resultant action is repulsion away from cell type C1. *Melanocytes can also differentiate randomly, independent of any other cell type.
- Supplementary file 6. Parameters implemented in the model.
- Transparent reporting form

Data availability

All mathematical modelling assumptions/ methods have been provided in the supplementary material. Code relating to this paper have been made available on github, at https://github.com/

JenniferOwen/Zebrafish-stripe-model (copy archived at https://github.com/elifesciences-publications/Zebrafish-stripe-model).

References

Asai R, Taguchi E, Kume Y, Saito M, Kondo S. 1999. Zebrafish leopard gene as a component of the putative reaction-diffusion system. *Mechanisms of Development* **89**:87–92. DOI: https://doi.org/10.1016/S0925-4773(99)00211-7, PMID: 10559483

Binder BJ, Simpson MJ. 2013. Quantifying spatial structure in experimental observations and agent-based simulations using pair-correlation functions. *Physical Review E* **88**:1–10. DOI: https://doi.org/10.1103/PhysRevE.88.022705

Bloomfield JM, Painter KJ, Sherratt JA. 2011. How does cellular contact affect differentiation mediated pattern formation? *Bulletin of Mathematical Biology* **73**:1529–1558. DOI: https://doi.org/10.1007/s11538-010-9578-4, PMID: 20798994

Budi EH, Patterson LB, Parichy DM. 2011. Post-embryonic nerve-associated precursors to adult pigment cells: genetic requirements and dynamics of morphogenesis and differentiation. *PLOS Genetics* **7**:e1002044. DOI: https://doi.org/10.1371/journal.pgen.1002044, PMID: 21625562

Bullara D, De Decker Y. 2015. Pigment cell movement is not required for generation of turing patterns in zebrafish skin. *Nature Communications* **6**:6971. DOI: https://doi.org/10.1038/ncomms7971, PMID: 25959141

Cal L, Suarez-Bregua P, Comesaña P, Owen J, Braasch I, Kelsh R, Cerdá-Reverter JM, Rotllant J. 2019. Countershading in zebrafish results from an Asip1 controlled dorsoventral gradient of pigment cell differentiation. *Scientific Reports* **9**:3449. DOI: https://doi.org/10.1038/s41598-019-40251-z, PMID: 30837630

Ceinos RM, Guillot R, Kelsh RN, Cerdá-Reverter JM, Rotllant J. 2015. Pigment patterns in adult fish result from superimposition of two largely independent pigmentation mechanisms. *Pigment Cell & Melanoma Research* **28**:196–209. DOI: https://doi.org/10.1111/pcmr.12335, PMID: 25469713

Clark E, Peel AD. 2018. Evidence for the temporal regulation of insect segmentation by a conserved sequence of transcription factors. *Development* **145**:dev.155580. DOI: https://doi.org/10.1242/dev.155580

Davies JA. 2002. Do different branching epithelia use a conserved developmental mechanism? *BioEssays* **24**:937–948. DOI: https://doi.org/10.1002/bies.10161, PMID: 12325126

Dini S, Binder BJ, Green JEF. 2018. Understanding interactions between populations: individual based modelling and quantification using pair correlation functions. *Journal of Theoretical Biology* **439**:50–64. DOI: https://doi.org/10.1016/j.jtbi.2017.11.014, PMID: 29197512

Dooley CM, Mongera A, Walderich B, Nüsslein-Volhard C. 2013. On the embryonic origin of adult melanophores: the role of ErbB and kit signalling in establishing melanophore stem cells in zebrafish. *Development* **140**:1003–1013. DOI: https://doi.org/10.1242/dev.087007, PMID: 23364329

Eom DS, Inoue S, Patterson LB, Gordon TN, Slingwine R, Kondo S, Watanabe M, Parichy DM. 2012. Melanophore migration and survival during zebrafish adult pigment stripe development require the immunoglobulin superfamily adhesion molecule Igsf11. *PLOS Genetics* **8**:e1002899. DOI: https://doi.org/10.1371/journal.pgen.1002899, PMID: 22916035

Eom DS, Bain EJ, Patterson LB, Grout ME, Parichy DM. 2015. Long-distance communication by specialized cellular projections during pigment pattern development and evolution. *eLife* **4**:e12401. DOI: https://doi.org/10.7554/eLife.12401, PMID: 26701906

Eom DS, Parichy DM. 2017. A macrophage relay for long-distance signaling during postembryonic tissue remodeling. *Science* **355**:1317–1320. DOI: https://doi.org/10.1126/science.aal2745, PMID: 28209639

Fadeev A, Krauss J, Frohnhöfer HG, Irion U, Nüsslein-Volhard C. 2015. Tight junction protein 1a regulates pigment cell organisation during zebrafish colour patterning. *eLife* **4**:e06545. DOI: https://doi.org/10.7554/eLife.06545, PMID: 25915619

Frohnhöfer HG, Krauss J, Maischein HM, Nüsslein-Volhard C. 2013. Iridophores and their interactions with other chromatophores are required for stripe formation in zebrafish. *Development* **140**:2997–3007. DOI: https://doi.org/10.1242/dev.096719, PMID: 23821036

Gavagnin E, Owen JP, Yates CA. 2018. Pair correlation functions for identifying spatial correlation in discrete domains. *Physical Review E* **97**:062104. DOI: https://doi.org/10.1103/PhysRevE.97.062104, PMID: 30011502

Gillespie DT. 1977. Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry* **81**:2340–2361. DOI: https://doi.org/10.1021/j100540a008

Hamada H, Watanabe M, Lau HE, Nishida T, Hasegawa T, Parichy DM, Kondo S. 2014. Involvement of Delta/Notch signaling in zebrafish adult pigment stripe patterning. *Development* **141**:318–324. DOI: https://doi.org/10.1242/dev.099804, PMID: 24306107

Hirata M, Nakamura K, Kanemaru T, Shibata Y, Kondo S. 2003. Pigment cell organization in the hypodermis of zebrafish. *Developmental Dynamics* **227**:497–503. DOI: https://doi.org/10.1002/dvdy.10334, PMID: 12889058

Hirata M, Nakamura K, Kondo S. 2005. Pigment cell distributions in different tissues of the zebrafish, with special reference to the striped pigment pattern. *Developmental Dynamics* **234**:293–300. DOI: https://doi.org/10.1002/dvdy.20513, PMID: 16110504

Inoue S, Kondo S, Parichy DM, Watanabe M. 2014. Tetraspanin 3c requirement for pigment cell interactions and boundary formation in zebrafish adult pigment stripes. *Pigment Cell & Melanoma Research* **27**:190–200. DOI: https://doi.org/10.1111/pcmr.12192

Irion U, Frohnhöfer HG, Krauss J, Çolak Champollion T, Maischein HM, Geiger-Rudolph S, Weiler C, Nüsslein-Volhard C. 2014. Gap junctions composed of connexins 41.8 and 39.4 are essential for colour pattern formation in zebrafish. *eLife* **3**:e05125. DOI: https://doi.org/10.7554/eLife.05125, PMID: 25535837

Iwashita M, Watanabe M, Ishii M, Chen T, Johnson SL, Kurachi Y, Okada N, Kondo S. 2006. Pigment pattern in Jaguar/obelix zebrafish is caused by a Kir7.1 mutation: implications for the regulation of melanosome movement. *PLOS Genetics* **2**:e197. DOI: https://doi.org/10.1371/journal.pgen.0020197, PMID: 17121467

Kelsh RN, Brand M, Jiang YJ, Heisenberg CP, Lin S, Haffter P, Odenthal J, Mullins MC, van Eeden FJ, Furutani-Seiki M, Granato M, Hammerschmidt M, Kane DA, Warga RM, Beuchle D, Vogelsang L, Nüsslein-Volhard C. 1996. P. zebrafish pigmentation mutations and the processes of neural crest development. *Development* **123**: 369–389. PMID: 9007256

Kelsh RN, Harris ML, Colanesi S, Erickson CA. 2009. Stripes and belly-spots – a review of pigment cell morphogenesis in vertebrates. *Seminars in Cell & Developmental Biology* **20**:90–104. DOI: https://doi.org/10.1016/j.semcdb.2008.10.001, PMID: 18977309

Kelsh RN, Sosa KC, Owen JP, Yates CA. 2017. Zebrafish adult pigment stem cells are multipotent and form pigment cells by a progressive fate restriction process: clonal analysis identifies shared origin of all pigment cell types. *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology* **39**:201600234. DOI: https://doi.org/10.1002/bies.201600234

Kondo S. 2017. An updated kernel-based turing model for studying the mechanisms of biological pattern formation. *Journal of Theoretical Biology* **414**:120–127. DOI: https://doi.org/10.1016/j.jtbi.2016.11.003, PMID: 27838459

Kondo S, Asal R. 1995. A reaction-diffusion wave on the skin of the marine angelfish Pomacanthus. *Nature* **376**: 765–768. DOI: https://doi.org/10.1038/376765a0, PMID: 24547605

Kondo S, Watanabe M. 2015. Black, yellow, or silver: which one leads skin pattern formation? *Pigment Cell & Melanoma Research* **28**:2–4. DOI: https://doi.org/10.1111/pcmr.12328, PMID: 25367546

Krauss J, Frohnhöfer HG, Walderich B, Maischein HM, Weiler C, Irion U, Nüsslein-Volhard C. 2014. Endothelin signalling in Iridophore development and stripe pattern formation of zebrafish. *Biology Open* **3**:503–509. DOI: https://doi.org/10.1242/bio.20148441, PMID: 24857848

Lister JA, Robertson CP, Lepage T, Johnson SL, Raible DW. 1999. Nacre encodes a zebrafish Microphthalmia-Related protein that regulates Neural-Crest-Derived pigment cell fate. *Development* **126**:3757–3767. PMID: 10433906

Lopes SS, Yang X, Müller J, Carney TJ, McAdow AR, Rauch GJ, Jacoby AS, Hurst LD, Delfino-Machín M, Haffter P, Geisler R, Johnson SL, Ward A, Kelsh RN. 2008. Leukocyte tyrosine kinase functions in pigment cell development. *PLOS Genetics* **4**:e1000026. DOI: https://doi.org/10.1371/journal.pgen.1000026, PMID: 18369445

Maderspacher F, Nüsslein-Volhard C. 2003. Formation of the adult pigment pattern in zebrafish requires leopard and Obelix dependent cell interactions. *Development* **130**:3447–3457. DOI: https://doi.org/10.1242/dev.00519, PMID: 12810592

Mahalwar P, Walderich B, Singh AP, Nüsslein-Volhard C. 2014. Local reorganization of xanthophores fine-tunes and colors the striped pattern of zebrafish. *Science* **345**:1362–1364. DOI: https://doi.org/10.1126/science.1254837, PMID: 25214630

Mahalwar P, Singh AP, Fadeev A, Nüsslein-Volhard C, Irion U. 2016. Heterotypic interactions regulate cell shape and density during color pattern formation in zebrafish. *Biology Open* **5**:1680–1690. DOI: https://doi.org/10.1242/bio.022251, PMID: 27742608

Maini PK, Woolley TE, Baker RE, Gaffney EA, Lee SS. 2012. Turing's model for biological pattern formation and the robustness problem. *Interface Focus* **2**:487–496. DOI: https://doi.org/10.1098/rsfs.2011.0113, PMID: 23919129

McMenamin SK, Bain EJ, McCann AE, Patterson LB, Eom DS, Waller ZP, Hamill JC, Kuhlman JA, Eisen JS, Parichy DM. 2014. Thyroid hormone-dependent adult pigment cell lineage and pattern in zebrafish. *Science* **345**:1358–1361. DOI: https://doi.org/10.1126/science.1256251, PMID: 25170046

Mellgren EM, Johnson SL. 2006. Pyewacket, a new zebrafish fin pigment pattern mutant. *Pigment Cell Research* **19**:232–238. DOI: https://doi.org/10.1111/j.1600-0749.2006.00311.x, PMID: 16704457

Metz HC, Manceau M, Hoekstra HE. 2011. Turing patterns: how the fish got its spots. *Pigment Cell & Melanoma Research* **24**:12–14. DOI: https://doi.org/10.1111/j.1755-148X.2010.00814.x, PMID: 21118391

Moreira J, Deutsch A. 2005. Pigment pattern formation in zebrafish during late larval stages: a model based on local interactions. *Developmental Dynamics* **232**:33–42. DOI: https://doi.org/10.1002/dvdy.20199, PMID: 15543601

Nakamasu A, Takahashi G, Kanbe A, Kondo S. 2009. Interactions between zebrafish pigment cells responsible for the generation of turing patterns. *PNAS* **106**:8429–8434. DOI: https://doi.org/10.1073/pnas.0808622106, PMID: 19433782

Painter KJ, Bloomfield JM, Sherratt JA, Gerisch A. 2015. A nonlocal model for contact attraction and repulsion in heterogeneous cell populations. *Bulletin of Mathematical Biology* **77**:1132–1165. DOI: https://doi.org/10.1007/s11538-015-0080-x, PMID: 25963245

Parichy DM, Mellgren EM, Rawls JF, Lopes SS, Kelsh RN, Johnson SL. 2000a. Mutational analysis of endothelin receptor b1 (rose) during neural crest and pigment pattern development in the zebrafish *Danio rerio*. *Developmental Biology* **227**:294–306. DOI: https://doi.org/10.1006/dbio.2000.9899, PMID: 11071756

Parichy DM, Ransom DG, Paw B, Zon LI, Johnson SL. 2000b. An orthologue of the kit-related gene fms is required for development of neural crest-derived xanthophores and a subpopulation of adult melanocytes in the zebrafish, *Danio rerio*. *Development* **127**:3031–3044. PMID: 10862741

Parichy DM. 2006. Evolution of Danio pigment pattern development. *Heredity* **97**:200–210. DOI: https://doi.org/10.1038/sj.hdy.6800867, PMID: 16835593

Parichy DM, Elizondo MR, Mills MG, Gordon TN, Engeszer RE. 2009. Normal table of postembryonic zebrafish development: staging by externally visible anatomy of the living fish. *Developmental Dynamics* **238**:2975–3015. DOI: https://doi.org/10.1002/dvdy.22113, PMID: 19891001

Parichy DM, Turner JM. 2003. Zebrafish Puma mutant decouples pigment pattern and somatic metamorphosis. *Developmental Biology* **256**:242–257. DOI: https://doi.org/10.1016/S0012-1606(03)00015-0, PMID: 12679100

Patterson LB, Bain EJ, Parichy DM. 2014. Pigment cell interactions and differential xanthophore recruitment underlying zebrafish stripe reiteration and Danio pattern evolution. *Nature Communications* **5**:5299. DOI: https://doi.org/10.1038/ncomms6299, PMID: 25374113

Patterson LB, Parichy DM. 2013. Interactions with iridophores and the tissue environment required for patterning melanophores and xanthophores during zebrafish adult pigment stripe formation. *PLOS Genetics* **9**: e1003561. DOI: https://doi.org/10.1371/journal.pgen.1003561, PMID: 23737760

Patterson LB, Parichy DM. 2019. Zebrafish pigment pattern formation: insights into the development and evolution of adult form. *Annual Review of Genetics* **53**:505–530. DOI: https://doi.org/10.1146/annurev-genet-112618-043741, PMID: 31509458

Singh AP, Schach U, Nüsslein-Volhard C. 2014. Proliferation, dispersal and patterned aggregation of iridophores in the skin prefigure striped colouration of zebrafish. *Nature Cell Biology* **16**:604–611. DOI: https://doi.org/10.1038/ncb2955, PMID: 24776884

Singh AP, Nüsslein-Volhard C. 2015. Zebrafish stripes as a model for vertebrate colour pattern formation. *Current Biology* **25**:R81–R92. DOI: https://doi.org/10.1016/j.cub.2014.11.013, PMID: 25602311

Svetic V, Hollway GE, Elworthy S, Chipperfield TR, Davison C, Adams RJ, Eisen JS, Ingham PW, Currie PD, Kelsh RN. 2007. Sdf1a patterns zebrafish melanophores and links the somite and melanophore pattern defects in choker mutants. *Development* **134**:1011–1022. DOI: https://doi.org/10.1242/dev.02789, PMID: 17267445

Takahashi G, Kondo S. 2008. Melanophores in the stripes of adult zebrafish do not have the nature to gather, but disperse when they have the space to move. *Pigment Cell & Melanoma Research* **21**:677–686. DOI: https://doi.org/10.1111/j.1755-148X.2008.00504.x, PMID: 19067972

Tickle C. 2006. Making digit patterns in the vertebrate limb. *Nature Reviews Molecular Cell Biology* **7**:45–53. DOI: https://doi.org/10.1038/nrm1830, PMID: 16493412

Volkening A, Sandstede B. 2015. Modelling stripe formation in zebrafish: an agent-based approach. *Journal of the Royal Society Interface* **12**:20150812. DOI: https://doi.org/10.1098/rsif.2015.0812, PMID: 26538560

Volkening A, Sandstede B. 2018. Iridophores as a source of robustness in zebrafish stripes and variability in Danio patterns. *Nature Communications* **9**:3231. DOI: https://doi.org/10.1038/s41467-018-05629-z, PMID: 30104716

Walderich B, Singh AP, Mahalwar P, Nüsslein-Volhard C. 2016. Homotypic cell competition regulates proliferation and tiling of zebrafish pigment cells during colour pattern formation. *Nature Communications* **7**: 11462. DOI: https://doi.org/10.1038/ncomms11462, PMID: 27118125

Watanabe M, Iwashita M, Ishii M, Kurachi Y, Kawakami A, Kondo S, Okada N. 2006. Spot pattern of *leopard Danio* is caused by mutation in the zebrafish *connexin41.8* gene. *EMBO Reports* **7**:893–897. DOI: https://doi.org/10.1038/sj.embor.7400757, PMID: 16845369

Watanabe M, Kondo S. 2015. Is pigment patterning in fish skin determined by the turing mechanism? *Trends in Genetics* **31**:88–96. DOI: https://doi.org/10.1016/j.tig.2014.11.005, PMID: 25544713

Yamaguchi M, Yoshimoto E, Kondo S. 2007. Pattern regulation in the stripe of zebrafish suggests an underlying dynamic and autonomous mechanism. *PNAS* **104**:4790–4793. DOI: https://doi.org/10.1073/pnas.0607790104, PMID: 17360399

Yamanaka H, Kondo S. 2014. In vitro analysis suggests that difference in cell movement during direct interaction can generate various pigment patterns in vivo. *PNAS* **111**:1867–1872. DOI: https://doi.org/10.1073/pnas.1315416111, PMID: 24449859

Appendix 1

Implementing the model

This text provides a full description of how a simulation of a WT fish is implemented. All notation used in appendix 1 is summarised in ***Supplementary files 1*** and ***2***.

Model overview

We model five different cell types; yellow xanthophores (X), unpigmented xanthoblasts (X^b), black melanocytes (M), silver dense S-iridophores (I^d) and blue loose (I^l) S-iridophores. We account for the three separate hypodermis layers upon which the cells lie as three separate lattice domains, a xanthophore layer represented by matrix $\boldsymbol{X}$, a melanocyte layer represented by matrix $\boldsymbol{M}$, and an S-iridophore layer represented by a matrix $\boldsymbol{I}$. We denote the length and height of domain $D \in \{\boldsymbol{X}, \boldsymbol{M}, \boldsymbol{I}\}$ at time t in terms of the numbers of lattice sites as $\Pi_D^H(t)$, $\Pi_D^L(t)$ respectively. At any time, t, a lattice site at row i and column j in $\boldsymbol{M}$ denoted $\boldsymbol{M}(i,j,t)$ can either be occupied by M, i.e. $\boldsymbol{M}(i,j,t) = M$ or be unoccupied, $\boldsymbol{M}(i,j,t) = 0$. Similarly $\boldsymbol{I}(i,j,t)$ can either be occupied by I^l, i.e. $\boldsymbol{I}(i,j,t) = I^l$, be occupied by I^d, i.e. $\boldsymbol{I}(i,j,t) = I^d$, or be unoccupied, $\boldsymbol{I}(i,j,t) = 0$. Finally, $\boldsymbol{X}(i,j,t)$ can either be occupied by X, i.e. $\boldsymbol{X}(i,j,t) = X$, X^b, i.e. $\boldsymbol{X}(i,j,t) = X^b$ or be unoccupied, $\boldsymbol{X}(i,j,t) = 0$. This captures volume exclusion rules on each lattice layer which require that at any given time, any lattice site on the domain can be occupied by at most one cell. If at any time during the simulation an action is chosen which would break this rule, this action is aborted. To account for the different packing densities of the cell types, lattice sites on $\boldsymbol{X}$ and $\boldsymbol{I}$ are of size 0.02 mm $\times$ 0.02 mm, whilst lattice sites in $\boldsymbol{M}$ are of size 0.04 mm $\times$ 0.04 mm (***Parichy et al., 2000a***). We denote these as $\Delta_{\boldsymbol{X}} = \Delta_{\boldsymbol{I}} = 0.02$ mm and $\Delta_{\boldsymbol{M}} = 0.04$ mm. Hence the simulated height and length at any given time t is given as $\Omega_H(t) = \Pi_D^L(t)\Delta_D$, $\Omega_L(t) = \Pi_D^H(t)\Delta_D$. An illustration of the three layers and their corresponding cell types are given in ***Figure 2A*** in the main text. The lattices correspond to matrices;

$$\boldsymbol{X} = \begin{bmatrix} 0 & 0 & X^b & 0 & 0 & X & 0 & 0 & X^b & 0 & X & 0 \\ 0 & 0 & 0 & X & 0 & 0 & 0 & 0 & X & 0 & 0 & 0 \\ 0 & X & 0 & X & 0 & 0 & X^b & 0 & 0 & 0 & 0 & 0 \\ X & 0 & X & 0 & X & X & 0 & X & 0 & X & X & X \\ 0 & 0 & X^b & 0 & X^b & 0 & 0 & 0 & 0 & X & 0 & 0 \\ 0 & X & 0 & 0 & 0 & 0 & X & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{I} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & I^l & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & I^l & 0 & 0 & 0 & I^l & I^l & 0 & 0 \\ I^d & I^d & I^d & I^d & I^d & I^d & I^d & I^d & I^d & I^d & I^d & I^d \\ I^d & I^d & I^d & I^d & I^d & I^d & I^d & I^d & I^d & I^d & I^d & I^d \\ 0 & 0 & I^l & 0 & I^l & I^l & I^l & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{M} = \begin{bmatrix} M & 0 & M & M & 0 & 0 \\ 0 & 0 & 0 & M & 0 & 0 \\ 0 & M & 0 & 0 & M & 0 \end{bmatrix}$$

where $\Pi_{\boldsymbol{X}}^H = \Pi_{\boldsymbol{I}}^H = 6$, $\Pi_{\boldsymbol{X}}^L = \Pi_{\boldsymbol{I}}^L = 12$, $\Pi_{\boldsymbol{M}}^H = 3$, $\Pi_{\boldsymbol{M}}^L = 6$.

Relating size to stage

Zebrafish development is described in the literature with respect to the standard length (SL) of the fish (***Parichy et al., 2009***). We associate our domain at time t to different zebrafish developmental stages by calculating a simulated SL (Ω_{SL}) from the size of the domain at time t and relating this to the development stage using ***Supplementary file 3***. Since growth is linear with respect to time (in real life and modelled as so) and we initialise the domain 5.7 mm

shorter than the real width (SL), the corresponding simulated SL for any simulation at time t can be calculated by:

$$\Omega_{SL}(t) = \Delta_D \times \Pi_D^L(t) + 5.7\text{mm}, \tag{2}$$

where $D \in \{\boldsymbol{X}, \boldsymbol{M}, \boldsymbol{I}\}$ and $\Pi_D^H(t)$ is the number of sites in the y direction of domain type D at time t. The simulated height $\Omega_H(t) = \Delta_D \times \Pi_D^H(t)$ directly corresponds to the height at the anterior margin of the anal fin (abbreviated as HAA) of the fish at all times.

Initial conditions (WT only)

The simulation is initialised to represent the zebrafish half way through stage PB at approximately 25dpf. At this stage the fish is approximately 1 mm in height (i.e., HAA is 1 mm) and 7.7 mm in length (i.e. SL is 7.7 mm). We model the full height and a subsection of the full length for convenience. We set $\Omega_H(0) = 1mm$ and $\Omega_L(0) = 2mm$. Therefore, $\Pi_{\boldsymbol{M}}^L(0)$=25, $\Pi_{\boldsymbol{M}}^L(0)$=50, $\Pi_{\boldsymbol{X}}^H(0) = \Pi_{\boldsymbol{I}}^H(0) = 50$ and $\Pi_{\boldsymbol{X}}^L(0) = \Pi_{\boldsymbol{I}}^L(0) = 100$.

The initial occupancies for a WT simulation comprise three rows of I^d along the middle of the fish, that is, sites on rows 24 to 26 on $\boldsymbol{I}$ are fully occupied with I^d cells at time $t = 0$. This represents the S-iridophores which differentiate along the horizontal myoseptum between 21 and 25 dpf. We also place $N_M^T(0) = 50$ uniformly at random on $\boldsymbol{M}$ so that the initial occupancy density of melanocytes is

$$\frac{N_M^T(0)}{\Pi_{\boldsymbol{M}}^L(0) \times \Pi_{\boldsymbol{M}}^H(0)} = 0.04, \tag{3}$$

where $N_C^T(t)$ for any cell type $C \in \{M, X, X^b, I^d, I^l\}$ denotes the total number of cells of that type on the domain at time t. These scattered melanocytes represent the initial larval pattern at low density dispersed across the domain at the beginning of pattern metamorphosis. Similarly, we occupy the domain so that the initial density of xanthoblasts is 0.4. This density is chosen from evidence that larval xanthophores, which de-differentiate around 5 dpf, proliferate and cover much of the entire domain at the start of metamorphosis (***Mahalwar et al., 2014***). We do this by placing $N_{X^b}(0) = 2000$ uniformly at random on $\boldsymbol{X}$ so

$$\frac{N_{X^b}^T(0)}{\Pi_{\boldsymbol{X}}^L(0) \times \Pi_{\boldsymbol{X}}^H(0)} = 0.4. \tag{4}$$

This is to represent the de-differentiated larval xanthophores which initially appear around 5 dpf, subsequently lose pigment (becoming xanthoblasts) and proliferate between 5 and 25 dpf. For all other cell types i.e. $C \in \{X, I^l\}$, $N_C^T(0)$ = 0.

Model iteration

The model is updated in continuous time according to the Gillespie algorithm (***Gillespie, 1977***) (see ***Figure 2B*** in the main text for an illustration). In our simulation we allow two different event types: continuous and fixed. A continuous event is an event that can happen at any point throughout the simulated time. For example; melanocyte birth or death. A fixed event is an event that occurs once during the simulation, and happens upon meeting predetermined conditions. In WT fish, we define only one fixed event. We assume there is appearance of metamorphic xanthophores X along the horizontal myoseptum at stage SP. This is justified by observations of *shd* in which delayed appearance of metamorphic X in interstripe X0 were noted. We found this delayed appearance was important for generating pseudo-stripe patterns in *rse* and *shd* which have reduced or entirely absent S-iridophores. We model this by occupying all the sites, regardless of prior occupancy, along the middle three rows of $\boldsymbol{X}$ with xanthophores (X) at stage SP that is sites $\boldsymbol{X}(i, j, t_{SP}) = X$ for all $i \in \lceil \frac{\Pi_X^H(t_{SP})}{2} \rceil - 1 : \lceil \frac{\Pi_X^H(t_{SP})}{2} \rceil + 1$ where t_{SP} denotes the time that the given simulation reaches the start of stage SP.

There are 15 continuous event types. These events are summarised in ***Supplementary file 4***. They comprise cell birth, death, movement and shape transitions as well as growth of the domain.

An overview of how the model is updated is given in ***Figure 2B*** in the main text and can be described as follows. At any given time t, the model is first assessed for meeting the criteria of a fixed event. If the model meets the criteria, the fixed event occurs, is subsequently marked as complete and the simulation continues. If no fixed time event is to be implemented then one of the fifteen possible continuous time events is attempted. First an exponentially distributed waiting time

$$\tau = \frac{1}{\alpha_0}\log\left(\frac{1}{u_0}\right), \tag{5}$$

is generated until the next continuous 'event' occurs where $\alpha_0 = \sum_{i=1}^{15}\alpha_i(t)$ is the sum of all of the event propensities given in ***Supplementary file 4*** and u_0 is a uniformly distributed random number in $(0,1)$ (i.e. $u_0 \sim U(0,1)$). Next, we generate $u_1 \sim U(0,1)$ to determine which event occurs at time $t+\tau$. The event i that satisfies

$$\sum_{j=1}^{i-1}\frac{\alpha_j(t)}{\alpha_0} \leq u_1 < \sum_{j=1}^{i}\frac{\alpha_j(t)}{\alpha_0} \tag{6}$$

is chosen to occur and the domain is updated accordingly (provided conditions required for that event to occur are met). Time is also updated. $t = t+\tau$. We continue this process iteratively, checking for fixed events, then subsequently generating a time for the next continuous event to occur. The process repeats until we reach the end of pattern metamorphosis, marked by length condition $\Omega_{SL} = 13.5mm$. The algorithm is stochastic in the sense that, within any given simulation, there is variance with regard to the exact rate and order of event occurrence, just as in real fish.

Modeling cell–cell interactions

Continuous events 1–15 given in ***Supplementary file 4*** are mediated by different cell-cell interactions. Cells interact on the fish skin at both short (neighbouring cells) and long (half a stripe width) range, possibly regulated by direct contact, dendrites, or longer extensions (filopodia or airinemes). In our model, uniform disks, with radii on the order of the distance 0.04 mm account for short-range interactions, and an annulus with an outer radius of approximately half an adult stripe width 0.24mm represent long-range dynamics.

We denote $S_D(r,k)$ and $L_D(r,k)$ where $D \in \{\boldsymbol{X},\boldsymbol{I},\boldsymbol{M}\}$ is the set of site positions (i,j) such that $D(i,j)$ is in the short (0.04 mm) or long (0.24 mm) distance respectively from a focal site at position $D(r,k)$. Note that $S_D(r,k)$ and $L_D(r,k)$ are both different for different domain types due to the different lattice site sizes. $S_D(r,k)$ and $L_D(r,k)$ are visualised for $D = \boldsymbol{M},\boldsymbol{X}$ in ***Figure 3A–H*** in the main text. Formulae for these sets are given in ***Supplementary file 2***.

To illustrate this see ***Figure 3A–H*** in the main text which compares the number of cells in the short (***Figure 3A–D***) and long (***Figure 3E–H***) range distance from a central site, $D(r,k)$, on different domain types. In this figure, M are represented as black circles. X are represented as yellow circles. ***Figure 2A*** is a visualisation of sites (marked in red) in $S_{\boldsymbol{M}}(r,k)$, where $\boldsymbol{M}(r,k)$ is the central site marked in grey. $S^{M}_{\boldsymbol{M}}(r,k,t)$ are the sites marked in red in which a melanocyte resides. The number of melanocytes in the short range distance from $\boldsymbol{M}(r,k)$ at time t is given by $N^{S}_{M,M}(r,k,t) = S^{M}_{\boldsymbol{M}}(r,k,t) = 2$. ***Figure 2B*** is a visualisation of the sites considered when calculating $N^{S}_{M,X}$ (formula given in ***Supplementary file 2***). In this example, $N^{S}_{M,X}(r,k,t) = 9$. To compare the number of M and X in the short range distance of $\boldsymbol{M}(r,k)$, we consider $wN^{S}_{M,M}(r,k,t) = 4\times 2 = 8$ which corresponds to the weighted value of melanocytes in the short range distance with $N^{S}_{M,X}(r,k,t) = 9$ the number of xanthophores in the short range. ***Figure 3C*** is a visualisation of sites (marked in red) in $S_{\boldsymbol{X}}(r,k)$, where $\boldsymbol{X}(r,k)$ is the central site marked in grey. $S^{X}_{\boldsymbol{X}}(r,k,t)$ are the sites marked in red within which a xanthophore resides. The number of

xanthophores in the short range distance from $\boldsymbol{X}(r,k)$ at time t is given by $N^S_{X,X}(r,k,t) = S^X_{\boldsymbol{X}}(r,k,t) = 7$. ***Figure 2D*** is a visualisation of the sites considered when calculating $N^S_{X,M}(r,k,t)$ (formula given in ***Supplementary file 2***). In this example, $N^S_{X,M}(r,k,t) = 14$. To compare the number of M and X in the short range distance of $\boldsymbol{X}(r,k)$, at time t, we would compare $N^S_{X,M}(r,k,t)$=14 with $N^S_{X,X}(r,k,t) = 7$. ***Figure 3E*** is a visualisation of sites (marked in red) in $L_{\boldsymbol{M}}(r,k)$, where $\boldsymbol{M}(r,k)$ is the central site marked in grey. $L^M_{\boldsymbol{M}}(r,k,t)$ are the sites marked in red in which a melanocyte resides. The number of melanocytes in the long range distance from $\boldsymbol{M}(r,k)$ is given by $N^L_{M,M}(r,k,t) = L^M_{\boldsymbol{M}}(r,k,t) = 20$. ***Figure 2F*** is a visualisation of the sites considered when calculating $N^L_{M,X}(r,k,t)$ (formula given in ***Supplementary file 2***). In this example, $N^L_{M,X}(r,k,t) = 16$. To compare the number of M and X in the long range distance of $\boldsymbol{M}(r,k)$, we consider $wN^L_{M,M}(r,k,t) = 4 \times 20 = 80$ which corresponds to the weighted value of melanocytes in the long range distance with $N^S_{M,X}(r,k,t) = 16$, the number of xanthophores in the long range. ***Figure 2G*** is a visualisation of sites (marked in red) in $L_{\boldsymbol{X}}(r,k)$, where $\boldsymbol{X}(r,k)$ is the central site marked in grey. $L^X_{\boldsymbol{X}}(r,k)$ are the sites marked in red in which a xanthophore resides. The number of xanthophores in the long range distance from $\boldsymbol{X}(r,k)$ is given by $N^L_{X,X} = L^X_{\boldsymbol{X}}(r,k,t) = 2$. ***Figure 3H*** is a visualisation of the sites considered when calculating $N^L_{X,M}(r,k,t)$ (formula given in ***Supplementary file 2***). In this example $N^L_{X,M}(r,k,t) = 20$. To compare the number of M and X in the long range distance of $\boldsymbol{X}(r,k)$, we consider $N^L_{X,M}(r,k,t) = 20$ with $N^L_{X,X}(r,k,t) = 2$.

We denote $S^C_D(r,k,t) \subset S_D(r,k)$ where C lies in layer D to be the sites in $S_D(r,k)$ occupied by cell type $C \in \{M, X, I^d, I^l, X^b\}$ at time t. We denote $N^S_{C_i,C_j}(r,k,t)$, $(N^L_{C_i,C_j}(r,k,t))$ as the number of cells of type C_j in the short (long) range distance 0.04 mm, (0.24 mm) from cell type C_i at $D_i(r,k)$. Hence the number of cells of type C_i in the short distance from another cell of the same type at $D(r,k)$ at time t is given by $N^S_{C_i,C_i}(r,k,t) = |S^{C_i}_D(r,k,t)|$. The formulae for $N^S_{C_i,C_j}(r,k,t)$, $N^L_{C_i,C_j}(r,k,t)$ where C_i does not equal C_j are more complicated and are given in ***Equations 7, 8*** (and ***Supplementary file 2*** for reference).

$$N^S_{C_1,C_2}(r,k,t) = \begin{cases} S^{C_2}_{D_2}(r,k,t) \text{if } D_1 = D_2 \text{ or } D_1 \text{and} D_2 \in \{\boldsymbol{X},\boldsymbol{I}\}. \\ \sum_{(i,j)\in S_{D_1}(r,k)} 1_{\boldsymbol{M}\left(\lceil \frac{i}{2}\rceil,\lceil \frac{j}{2}\rceil\right)=M} \text{if} D_1 \neq \boldsymbol{M},\, D_2 = \boldsymbol{M}. \\ \sum_{(i,j)\in S_{\boldsymbol{M}}(r,k)} 1_{D_2(2i,2j)=C_2} + 1_{D_2(2i-1,2j)=C_2} + 1_{D_2(2i,2j-1)=C_2} + \ldots \\ 1_{D_2(2i-1,2j-1)=C_2} \text{if } D_1 = \boldsymbol{M},\, D_2 \neq \boldsymbol{M}. \end{cases} \tag{7}$$

$$N^L_{C_1,C_2}(r,k,t) = \begin{cases} L^{C_2}_{D_2}(r,k,t) \text{if } D_1 = D_2 \text{ or } D_1 \text{and} D_2 \in \{\boldsymbol{X},\boldsymbol{I}\}. \\ \sum_{(i,j)\in L_{D_1}(r,k)} 1_{\boldsymbol{M}\left(\lceil \frac{i}{2}\rceil,\lceil \frac{j}{2}\rceil\right)=M} \text{if} D_1 \neq \boldsymbol{M},\, D_2 = \boldsymbol{M}. \\ \sum_{(i,j)\in L_{\boldsymbol{M}}(r,k)} 1_{D_2(2i,2j)=C_2} + 1_{D_2(2i-1,2j)=C_2} + 1_{D_2(2i,2j-1)=C_2} + \ldots \\ 1_{D_2(2i-1,2j-1)=C_2} \text{if } D_1 = \boldsymbol{M},\, D_2 \neq \boldsymbol{M}. \end{cases} \tag{8}$$

Simply, where domain types D do not have the same lattice site size Δ_D, the focal site coordinates (i,j) of the local neighbourhood undergo a transformation for the sites on a domain with a different size. For example, each site $\boldsymbol{X}(i,j)$, corresponds to a quarter of site $\boldsymbol{M}(\lceil \frac{i}{2}\rceil, \lceil \frac{j}{2}\rceil)$, this explains case two in ***Equation 7***. Similarly each site $\boldsymbol{M}(i,j)$ corresponds to four sites on $\boldsymbol{X}$, specifically $\boldsymbol{X}(2i,2j)$, $\boldsymbol{X}(2i,2j-1)$, $\boldsymbol{X}(2i-1,2j)$ and $\boldsymbol{X}(2i-1,2j-1)$. This explains case three in ***Equations 7 and 8***. Note that since M is four times larger than all other cells in our simulation, we provide a weighting system when comparing M with CM in some cases.

Boundary conditions

Boundary conditions are periodic along the horizontal boundaries and reflecting across the vertical boundaries. We implement periodic boundary conditions along the horizontal axis

based on the assumption that the rate at which cells leave along this axis is approximately equal to the rate at which cells enter the domain at the opposite side.

Continuous time events

In this text, we will describe how each of the fifteen events are simulated upon being selected to occur by the Gillespie algorithm. An overview of all continuous time events and their corresponding rates is given in ***Supplementary file 4***.

Movement (continuous time events 1–5)

We implement cell movement so that cells are biased towards cell types they are attracted to and away from cell types they are repelled by. The direction of the cell's movement is determined using an on-lattice attraction-repulsion mechanism based on a model described by ***Dini et al., 2018*** and is detailed as follows. If a cell is chosen to move, it is able to move in one of eight different possible orientations denoted by $O \in \boldsymbol{O} = \{No, So, Ea, We, NE, SE, SW, NW\}$. (See ***Figure 3(D)*** in the main text for an illustration of these directions).

The directional neighbours of each cell is given in ***Figure 2E,F***. ***Figure 2E*** demonstrates the directional neighbourhoods of a focal cell $C \in \{X, X^b, I^d, I^l\}$ located in position $D(i,j)$ (marked in grey) where $D \in \{\boldsymbol{X}, \boldsymbol{I}\}$. For the cell moving one space on D, the cell considers the occupancy of sites in $S_{k\Delta_X}^{D}(i,j)$ marked in red for $D = \boldsymbol{M}$ (top) and $D \in \{\boldsymbol{X}, \boldsymbol{I}\}$ (bottom). (Q–X) ***Figure 2F*** demonstrates the directional neighbourhoods of a focal cell M located in position $\boldsymbol{M}(i,j)$ (marked in grey). For the cell M moving one space on $\boldsymbol{M}$, the cell considers the occupancy of sites in $S_{k\Delta_M}^{D}(i,j)$ marked in red for $D \in \{\boldsymbol{X}, \boldsymbol{I}\}$ (top) and $D = \boldsymbol{M}$ (bottom).

We define the probability that a cell attempts to move in a given direction with orientation O as P_O (where $\sum_{O \in \boldsymbol{O}} P_O = 1$). We calculate P_O, using bias vector,

$$\underline{v} = \boldsymbol{A}\underline{a} + \boldsymbol{R}\underline{r}, \tag{9}$$

where $\boldsymbol{A}$ ($\boldsymbol{R}$) is a matrix whose entries are the number of neighbouring cells within different segments of the attraction (repulsion) range and $\underline{a}$ ($\underline{r}$) is a weight of attraction (repulsion) vector.

Matrices $\boldsymbol{A}$ and $\boldsymbol{R}$ are defined as follows.

$$\boldsymbol{A} = \begin{bmatrix} Ea_{C,C_1}(r,k,t) & Ea_{C,C_2}(r,k,t) & Ea_{C,C_3}(r,k,t) \\ We_{C,C_1}(r,k,t) & We_{C,C_2}(r,k,t) & W_{C,C_3}(r,k,t) \\ No_{C,C_1}(r,k,t) & No_{C,C_2}(r,k,t) & No_{C,C_3}(r,k,t) \\ So_{C,C_1}(r,k,t) & So_{C,C_2}(r,k,t) & So_{C,C_3}(r,k,t) \\ NW_{C,C_1}(r,k,t) & NW_{C,C_2}(r,k,t) & NW_{C,C_3}(r,k,t) \\ NE_{C,C_1}(r,k,t) & NE_{C,C_2}(r,k,t) & NE_{C,C_3}(r,k,t) \\ SW_{C,C_1}(r,k,t) & SW_{C,C_2}(r,k,t) & NE_{C,C_3}(r,k,t) \\ SE_{C,C_1}(r,k,t) & SE_{C,C_2}(r,k,t) & SE_{C,C_3}(r,k,t) \end{bmatrix}, \tag{10}$$

$$\boldsymbol{R} = \begin{bmatrix} We_{C,C_1}(r,k,t) & We_{C,C_2}(r,k,t) & We_{C,C_3}(r,k,t) \\ Ea_{C,C_1}(r,k,t) & Ea_{C,C_2}(r,k,t) & Ea_{C,C_3}(r,k,t) \\ So_{C,C_1}(r,k,t) & So_{C,C_2}(r,k,t) & So_{C,C_3}(r,k,t) \\ No_{C,C_1}(r,k,t) & No_{C,C_2}(r,k,t) & No_{C,C_3}(r,k,t) \\ NE_{C,C_1}(r,k,t) & NE_{C,C_2}(r,k,t) & NE_{C,C_3}(r,k,t) \\ NW_{C,C_1}(r,k,t) & NW_{C,C_2}(r,k,t) & NW_{C,C_3}(r,k,t) \\ SE_{C,C_1}(r,k,t) & SE_{C,C_2}(r,k,t) & SE_{C,C_3}(r,k,t) \\ SW_{C,C_1}(r,k,t) & SW_{C,C_2}(r,k,t) & NE_{C,C_3}(r,k,t) \end{bmatrix} \tag{11}$$

where for example, $O_{C,C_1}(r,k,t)$ denotes the number of cells of type $C_1 \in \{X, M, I^d, I^l, X^b\}$ within

range 0.04 mm, that are orientation O from the focal cell C located in position $D(r,k)$ at time t. This number is calculated differently depending on whether the focal site lies on lattice domain $D \in \{\boldsymbol{M},\boldsymbol{X},\boldsymbol{I}\}$. To compute $O_{C,C_1}(r,k,t)$ we calculate the following. For any O we define subsets $S_D^O(r,k,t)$ of $S_D(r,k,t)$ (where cell type C lies on domain type D) for each direction O;

$$S_D^{No}(r,k,t) = \{(i,j) \in S_D(r,k,t) \,|\, j>k\}, \tag{12}$$

$$S_D^{So}(r,k,t) = \{(i,j) \in S_D(r,k,t) \,|\, j<k\}, \tag{13}$$

$$\begin{aligned} S_D^{E}(r,k,t) = \; & \{(i,j) \in S_D(r,k,t) \,|\, (|r-i|<\Pi_D^L(t) - |r-i| \text{and} i>r) \\ & \text{or}(|r-i|>\Pi_D^L(t) - |r-i| \text{and} i<r)\}, \end{aligned} \tag{14}$$

$$\begin{aligned} S_D^{W}(r,k,t) = \; & \{(i,j) \in S_D(r,k,t) \,|\, (|r-i|<\Pi_D^L(t) - |r-i| \text{and} i<r) \\ & \text{or}(|r-i|>\Pi_D^L(t) - |r-i| \text{and} i>r)\}. \end{aligned} \tag{15}$$

These are the sets of sites in the short range distance, north, south, east and west of focal site $D(r,k)$. Examples for $D=\boldsymbol{M}$ and $\boldsymbol{X}$ are sites marked in red on the left of ***Figure 2E–F*** respectively. By extension we define

$$S_D^{NE}(r,k,t) = \{(i,j) \in S_D(r,k,t) \,|\, j \geq k \text{and}(|r-i|<\Pi_D^L(t) - |r-i| \text{and} i \geq r) \tag{16}$$

$$\text{or}(|r-i|>\Pi_D^L(t) - |r-i| \text{and} i \leq r)\}, \tag{17}$$

$$S_D^{NW}(r,k,t) = \{(i,j) \in S_D(r,k,t) \,|\, j \geq k \text{and}(|r-i|<\Pi_D^L(t) - |r-i| \text{and} i \leq r) \tag{18}$$

$$\text{or}(|r-i|>\Pi_D^L(t) - |r-i| \text{and} i \geq r)\}, \tag{19}$$

$$S_D^{SE}(r,k,t) = \{(i,j) \in S_D(r,k,t) \,|\, j \leq k \text{and}(|r-i|<\Pi_D^L(t) - |r-i| \text{and} i \geq r) \tag{20}$$

$$\text{or}(|r-i|>\Pi_D^L(t) - |r-i| \text{and} i \leq r)\}, \tag{21}$$

$$S_D^{SW}(r,k,t) = \{(i,j) \in S_D(r,k,t) \,|\, j \leq k \text{and}(|r-i|<\Pi_D^L(t) - |r-i| \text{and} i \leq r) \tag{22}$$

$$\text{or}(|r-i|>\Pi_D^L(t) - |r-i| \text{and} i \geq r)\}. \tag{23}$$

These are the sets of sites in the short range distance, north-east, north-west, south-east and south-west of focal site $D(r,k)$. We denote $S_D^{C,O}(r,k,t) \subseteq S_D^O(r,k,t)$ to be all the sites in short-range distance in orientation $O \in \{No,So,Ea,We,NE,SE,SW,NW\}$ occupied by cell type C. Hence we define

$$S_D^{C,O}(r,k,t) = \{(i,j) \in S_D^O(r,k,t) \,|\, D(i,j) = C\}. \tag{24}$$

Therefore to compute the number of cells in each of the eight directions, we compute where C lies on D and C_1 lies on D_1 by;

$$O_{C,C_1}(r,k,t) = \begin{cases} |S_{D_1}^{C,O}(r,k,t)| \text{where} D = D_1 \text{or} D, D_1 \in \{\boldsymbol{X},\boldsymbol{I}\}. \\ \sum\limits_{S_{D_1}^O(r,k)} \mathbb{1}_{\boldsymbol{M}\left(\lceil \frac{r}{2} \rceil, \lceil \frac{k}{2} \rfloor\right)=M} \text{where} D = M, D_1 \neq \boldsymbol{M}. \\ \sum\limits_{S_{\boldsymbol{M}}^O(r,k)} \mathbb{1}_{D(2r,2k)=C} + \mathbb{1}_{D(2r-1,2k)=C} \\ \quad + \mathbb{1}_{D(2r,2k-1)=C} + \mathbb{1}_{D(2r-1,2k-1)=C} \text{where} D \neq \boldsymbol{M}, D_1 = \boldsymbol{M}. \end{cases} \tag{25}$$

Next we define the weight of the attraction-repulsion vector by:

$$\underline{a} = \begin{bmatrix} a_{CC_1} \\ a_{CC_2} \\ a_{CC_3} \end{bmatrix}, \underline{r} = \begin{bmatrix} \rho_{CC_1} \\ \rho_{CC_2} \\ \rho_{CC_3} \end{bmatrix} \tag{26}$$

where $a_{CC_1}, a_{CC_2}, a_{CC_3}$ ($\rho_{CC_1}, \rho_{CC_2}, \rho_{CC_3}$) are the weights of attraction (repulsion) of cells of type C to cells of the type C_1, C_2 and C_3, respectively defined in ***Supplementary file 5***. Vectors $\underline{a}$ and $\underline{b}$ are the vectors used to calculate $\underline{v}$ in ***Equation 9***. To account for the difference in distance between adjacent neighbours and diagonal neighbours we normalise accordingly so that movement diagonally is proportionately less frequent than moving to adjacent squares.

$$\hat{P}_W = \frac{2v_1}{3\sum_{q=1}^{4} v_q}, \tag{27}$$

$$\hat{P}_E = \frac{2v_2}{3\sum_{q=1}^{4} v_q}, \tag{28}$$

$$\hat{P}_{So} = \frac{2v_3}{3\sum_{q=1}^{4} v_q}, \tag{29}$$

$$\hat{P}_{No} = \frac{2v_4}{3\sum_{q=1}^{4} v_q}, \tag{30}$$

$$\hat{P}_{NE} = \frac{v_5}{3\sum_{q=5}^{8} v_q}, \tag{31}$$

$$\hat{P}_{NW} = \frac{v_6}{3\sum_{q=5}^{8} v_q}, \tag{32}$$

$$\hat{P}_{SE} = \frac{v_7}{3\sum_{q=5}^{8} v_q}, \tag{33}$$

$$\hat{P}_{SW} = \frac{v_8}{3\sum_{q=5}^{8} v_q}. \tag{34}$$

Finally we normalise P_O s.t.;

$$P_O = \frac{\hat{P}_O}{\sum_{O \in \boldsymbol{O}} \hat{P}_O}. \tag{35}$$

To summarise, we simulate movement as follows. Given the event is chosen such that a cell of type C is allocated to move. We choose a random cell of type C. We call its position $D(r,k)$. Next values P_O are calculated using ***Equation 27–33*** and ***Equation 35***. A random number $u \sim U(0,1)$ is generated. The cell attempts movement to $D(r,k-1)$ (West) if $u \in [0, P_W)$, $D(r,k+1)$ (East) if $u \in [P_W, P_W + P_E)$ etc. up to $[\sum_{O \in \boldsymbol{O}} P_O - P_{SW}, \sum_{O \in \boldsymbol{O}} P_O]$. The cell movement is successful and the cell moves from site $D(r,k)$ to the nearest neighbouring site in direction O, provided this site is empty. Otherwise the movement is aborted. Movement bias is specified by weights of attraction and repulsion ($\underline{a}$ and $\underline{r}$). Relevant parameters for $a_{C_iC_j}$ and $\rho_{C_iC_j}$ for different C_i and C_j are given in ***Supplementary file 5*** . If no such interaction bias is specified then there is no known attraction or repulsion dynamics between those cell types in the literature and so we assume that movement is not biased by this cell type. Note that for cell X^b there are no known short range interactions between X^b and other cell types so we model X^b movement as random, i.e, $P_O = \frac{1}{8}$, for $O \in \boldsymbol{O}$ at all times.

S-iridophore, xanthophore and xanthoblast proliferation (continuous time events 6–7)

Given the event is chosen such that a cell of type $C \in \{X, X^b, I^d, I^l\}$ is determined to proliferate we choose a random cell of type C whose position is given by $D(r,k)$, $D \in \{\boldsymbol{X}, \boldsymbol{I}\}$. Next a random number $u \sim U(0,1)$ is generated. This number is used to determine a neighbouring site into which a mother cell can place a daughter cell. This site is $D(r,k+1)$ if $u \in [0, 1/4)$, $D(r+1,k)$ if $u \in [1/4, 1/2)$, $D(r-1,k)$ if $u \in [1/2, 3/4)$ or $D(r,k-1)$ if $u \in [3/4, 1]$.

For cell types $C \in \{I^d, I^l, X^b\}$, a proliferation event is successful if the site chosen is empty, and a new cell of the same type is placed into the chosen site, otherwise the event is aborted. In the case $C = X$, a proliferation event is successful if the site chosen for the daughter cell is unoccupied and

$$N^S_{X,X}(r,k,t) + N^S_{X,I^d}(r,k,t) > N^S_{X,I^l}(r,k,t) + N^S_{X,M}(r,k,t). \tag{36}$$

Equation 36 is based on the following assumptions: (1) Dense S-iridophores promote xanthoblasts differentiation into xanthophores in the short range (***Patterson and Parichy, 2013***). Dense S-iridophores express xanthogenic Colony Stimulating Factor-1 (Csf1) (***Patterson and Parichy, 2013***) which is essential for xanthophore differentiation, proliferation and survival, allowing unpigmented xanthoblasts near to the dense S-iridophores to mature into xanthophores (***Frohnhöfer et al., 2013***; ***Walderich et al., 2016***); (2) melanocytes inhibit xanthophore specification in the short range (***Nakamasu et al., 2009***).

Melanocyte differentiation (continuous time event 8)

If a melanocyte differentiation event is specified then a site $\boldsymbol{M}(r,k)$ is chosen at random. A differentiation event is successful and a new M appears in this site if the site $\boldsymbol{M}(r,k)$ is empty and the following is true:

$$N^L_{M,X}(r,k,t) + N^L_{M,I^d}(r,k,t) > w\alpha N^L_{M,M}(r,k,t) + \beta \tag{37}$$

$$\text{and } N^S_{M,X}(r,k,t) \leq \gamma w N^S_{M,M}(r,k,t), \tag{38}$$

$$\text{and } N^L_{M,I^d}(r,k,t) \leq \kappa, \tag{39}$$

where $\alpha = 2.5$, $\beta = \gamma = \kappa = 3, w = 4$. ***Equation 37*** is based on the findings that dense S-iridophores and xanthophores promote the differentiation of melanocytes in the long range. This conclusion is drawn from observations in ablation experiments (***Nakamasu et al., 2009***) and in the *pfe* mutant, which retains a high number of M(***Frohnhöfer et al., 2013***). ***Equation 38*** is based on observations that melanocytes and xanthophores compete in the short range (***Nakamasu et al., 2009***). ***Equation 39*** is based on observations that in WT fish, melanocyte centers rarely overlap with dense S-iridophores; however, melanocytes frequently settle adjacent to dense S-iridophores, suggesting short range inhibition (***Patterson and Parichy, 2013***).

We assume there is some melanocyte differentiation into empty space that is independent of cues from other cells. This is from observations of double mutant fish *nac;pfe* that do not produce S-iridophores or xanthophores but phenotypically display uniformly distributed melanocytes at adulthood (***Frohnhöfer et al., 2013***). Therefore, alternatively we also allow successful M differentiation if we generate a random number $u \in U(0,1)$ and we find that $u < 0.01$ in combination with the condition

$$N^S_{M,X}(r,k,t) + N^S_{M,I^d}(r,k,t) + N^S_{M,M}(r,k,t) = 0. \tag{40}$$

Within each attempt, criterion (***Equation 37–39***) is tried first. If this is not successful, criteria (***Equation 40***) is tested instead.

Xanthoblast differentiation (continuous time event 9)

If a xanthoblast differentiation event is chosen then an X^b is chosen at random from $\boldsymbol{X}$. Suppose the chosen X^b lies in site $\boldsymbol{X}(r,k)$. The differentiation event is successful and a X replaces the X^b in this site if the site $\boldsymbol{M}(\lceil\frac{r}{2}\rceil,\lceil\frac{k}{2}\rceil)$ is not occupied by a cell M (as melanocytes and xanthophores are known to compete in the short range [***Nakamasu et al., 2009***]) and if the following is true.

$$N^S_{X,X}(r,k,t)+N^S_{X,I^d}(r,k,t)>N^S_{X,M}(r,k,t)+N^S_{X,I^l}(r,k,t). \tag{41}$$

Equation 41 is based on the following assumptions; (1) Dense S-iridophores promote xanthoblasts differentiation into xanthophores in the short range (***Patterson and Parichy, 2013***). Dense S-iridophores express Csf1, allowing unpigmented xanthoblasts near to the dense S-iridophores to mature into xanthophores (***Frohnhöfer et al., 2013***; ***Walderich et al., 2016***). (2) melanocytes inhibit xanthophore specification in the short range (***Nakamasu et al., 2009***).

Alternatively the differentiation event is successful if:

$$N^S_{X,X}(r,k,t)+N^S_{X,M}(r,k,t)=0, \tag{42}$$

holds and a randomly generated number $u\in U(0,1)$ is such that u<0.01, or ***Equation(42)*** holds and either

$$t>31\,\mathrm{dpf}\,\mathrm{and}\frac{\mathrm{N^T_M(t)+N^T_X(t)+N^T_{I^d}(t)}}{\Pi^{\mathrm{L}}_{\mathrm{X}}(\mathrm{t})\times\Pi^{\mathrm{H}}_{\boldsymbol{X}}(\mathrm{t})}<0.2, \tag{43}$$

$$t>51\,\mathrm{dpf}\,\mathrm{and}\frac{\mathrm{N^T_M(t)+N^T_X(t)+N^T_{I^d}(t)}}{\Pi^{\mathrm{L}}_{\mathrm{X}}(\mathrm{t})\times\Pi^{\mathrm{H}}_{\boldsymbol{X}}(\mathrm{t})}<0.4, \tag{44}$$

is true i.e, the total cell density with all cells combined is less than either 0.2 or 0.4 for different time milestones. We enforce a lower probability (u < 0.01) for the alternative event (***Equation 42***) since this differentiation event is not influenced by cues from other cells and thus is less likely. ***Equation 43*** and ***Equation 44*** are based on assumptions that if the domain is empty after some key time points, this promotes delayed differentiation of xanthoblasts as is observed in *nac* (***Frohnhöfer et al., 2013***). Since melanocytes compete in the short range with xanthophores (***Nakamasu et al., 2009***) we only allow this to occur when constraints described by ***Equation(42)*** are simultaneously held, for consistency.

Iridophore transitions (continuous time event 10–11)

If an I^d to I^l shape transition is chosen, then an I^d is chosen uniformly at random from domain $\boldsymbol{I}$ and evaluated for meeting transition criteria. The transition is successful and the chosen I^d in position $\boldsymbol{I}(r,k)$ is replaced with I^l in this site if either both

$$N^L_{I^d,X}(r,k,t)>L_x \text{ and } N^S_{I^d,X}(r,k,t)<S_x, \tag{45}$$

or alternatively

$$N^S_{M,M}(r,k,t)\geq S_m, \tag{46}$$

where L_x,S_x,S_m=12, 9, one respectively. Alternatively if an I^l to I^d shape transition is chosen, then an I^l is chosen at random from domain $\boldsymbol{I}$ and evaluated for meeting transition criteria. This transition is successful and the chosen I^l in position $\boldsymbol{I}(r,k)$ is replaced with I^d in this site if either

$$N^L_{I^l,X}(r,k,t)<L_{x2} \text{ and } N^S_{I^l,M}(r,k,t)<S_{m2}, \tag{47}$$

or alternatively

$$N^S_{I^l,X}(r,k,t)>S_{x2} \text{ and } N^S_{I^l,M}(r,k,t)<S_{m2}, \quad (48)$$

where $L_{x2}=16$, $S_{x2}=4$ and $S_{m2}=1$. These conditions are based on observations of the induction set (**Frohnhöfer et al., 2013**) (for more details see Section "Modelling assumptions" in the main text). The parameters L_x, S_x, S_m, L_{x2}, S_{m2}, S_{x2} were chosen to give straight stripes with few breaks at stage J+ in WT simulations. However, with some small variations to these parameters,the patterns generated are qualitatively similar.

Melanocyte death (continuous time event 12)

If a melanocyte death event is chosen, then a cell of type M is chosen at random from $\boldsymbol{M}$. Suppose the chosen M lies in site $\boldsymbol{M}(r,k)$. The death event is successful and the melanocyte is removed from site $\boldsymbol{M}(r,k)$ i.e. $\boldsymbol{M}(r,k)$ is set to 0, if

$$wN^S_{M,M}(r,k,t)<N^S_{M,X}(r,k,t). \quad (49)$$

Equation 49 is based on findings that xanthophores and melanocytes compete in the short range (**Nakamasu et al., 2009**). Alternatively, the melanocyte death is successful if a randomly generated number $u\in U(0,1)$, is such that $u<0.001$ and both of the following two cases hold:

$$\sigma N^L_{M,X}(r,k,t) \leq N^L_{M,M}(r,k,t), \quad (50)$$

$$N^S_{M,I^l}(r,k,t) < \omega, \quad (51)$$

where $\sigma=\omega=3$. **Equation 50** is based on findings that M appear to inhibit the survival of M in the long range and that X promote the survival of M in the long range (**Takahashi and Kondo, 2008**). In **Equation 51**, I^l are a proxy for L-iridophores, which promote the survival of M in the short range (**Frohnhöfer et al., 2013**). We found this equation was important for maintaining melanocytes in mutant *pfe* in our simulations. Moreover, the two equations **Equation 50** and **Equation 51** combined were important for maintaining melanocytes in the interstripes in *pfe* (as seen in real fish) in our simulations. We found it was important that this alternative event had a low probability of success (0.001) but not zero. In *pfe* there are no xanthophores so **Equation 50** enforces melanocyte death where the number of melanocytes in the long range was greater than zero. Therefore, when there was too high a probability of success, in *pfe* simulations the stripe, interstripe structure that is usually maintained between WT and *pfe* was broken into randomly dispersed spots of melanocytes distance 0.24 mm apart. On the other hand, when it was too low, we did not observe melanocytes in the interstripe in our *pfe* simulations, unlike in real *pfe*.

Xanthoblast pulling event (continuous time event 13)

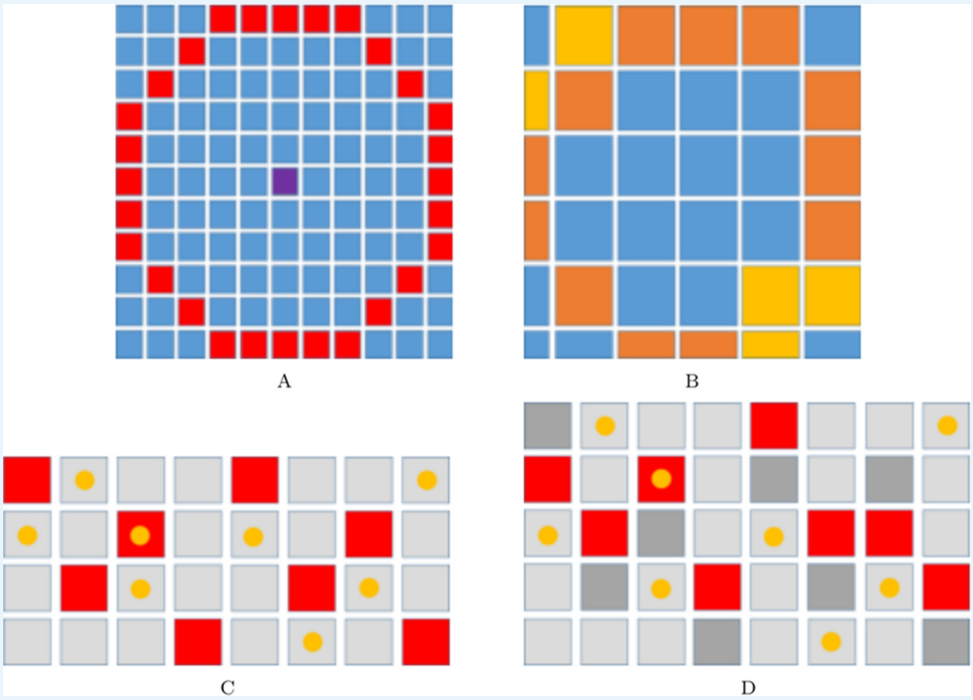
Suppose a 'xanthoblast pull melanocyte' event is chosen at time t, then we choose an X^b uniformly at random from all X^b occupying $\boldsymbol{X}$ that meets the following criterion: suppose the chosen X^b resides in site $\boldsymbol{X}(r,k)$, then the corresponding site on $\boldsymbol{M}$, $\boldsymbol{M}(\lceil\frac{r}{2}\rceil,\lceil\frac{k}{2}\rceil)$ must be empty. We simulate this by randomly choosing X^b on $\boldsymbol{X}$, and checking whether $\boldsymbol{M}(\lceil\frac{r}{2}\rceil,\lceil\frac{k}{2}\rceil)$ is empty. If so, we continue to the next stage. Otherwise we repeat through all X^b on $\boldsymbol{X}$ until we either find a suitable X^b. If no X^b satisfy this criteria at time t then we abort the cell-cell pulling event. Given we find a suitable X^b, the chosen X^b will attempt to attach and pull a melanocyte within range (airinemes extend to a length of up to 5–6 cell diameters away or 0.1–0.12 mm [**Eom et al., 2015**]) to position $\boldsymbol{M}(\lceil\frac{r}{2}\rceil,\lceil\frac{k}{2}\rceil)$. We simulate this as follows. First, we generate a random position (i,j) uniformly at random from the 28 possible lattice positions euclidean distance 0.1 mm from the site $X(r,k)$ given by P. These sites are shown in **Figure 1A–B** and their coordinates are given by;

$$P=\{(r+5,k),(r-5,k),(r,k+5),(r,k-5) \quad (52)$$
$$(r+1,k+5),(r-1,k+5),(r+1,k-5),(r-1,k-5), \quad (53)$$
$$(r+2,k+5),(r-2,k+5),(r+2,k-5),(r-2,k-5), \quad (54)$$
$$(r-3,k-4),(r+3,k+4),(r-3,k+4),(r+3,k-4), \quad (55)$$
$$(r-4,k-3),(r+4,k+3),(r-4,k+3),(r+4,k-3), \quad (56)$$
$$(r+5,k+1),(r-5,k+1),(r+5,k-1),(r-5,k-1), \quad (57)$$
$$(r+5,k+2),(r-5,k+2),(r+5,k-2),(r-5,k-2)\}. \quad (58)$$

Next we translate position (i,j) on $\boldsymbol{X}$ to its corresponding position on $\boldsymbol{M}$. If $\boldsymbol{M}(\lceil\frac{i}{2}\rceil,\lceil\frac{j}{2}\rceil,t)=M$, then the melanocyte is pulled from its site $\boldsymbol{M}(\lceil\frac{i}{2}\rceil,\lceil\frac{j}{2}\rceil)$ into $\boldsymbol{M}(\lceil\frac{r}{2}\rceil,\lceil\frac{k}{2}\rceil)i.e.\boldsymbol{M}(\lceil\frac{i}{2}\rceil,\lceil\frac{j}{2}\rceil,t)=0$ and $\boldsymbol{M}(\lceil\frac{r}{2}\rceil,\lceil\frac{k}{2}\rceil,t)=M$. The action is deemed complete. Otherwise, if $\boldsymbol{M}(i_M,j_M,t)=0$, i.e, the site is empty, then the process repeats for the same X^b. From here, P becomes $P\backslash(i,j)$ and a new (i,j) pair are generated uniformly at random from P until either an M is found within the range specified or until P is empty (i.e. after 28 tries). At this point, we will have determined that there were no melanocytes close enough to the xanthoblast for the xanthoblast to successfully pull.

Growth (continuous time events 14–15)

Given a growth event is chosen in the horizontal (vertical) direction, from each column (row), a site location is chosen uniformly at random at which insertion of new empty site above or below (randomly chosen) will occur. See ***Appendix 1—figure 1C-D*** for an example of a growth event in the vertical direction.



Appendix 1—figure 1. Implementation of xanthoblast-melanocyte communication and domain growth. (**A–B**) Range for xanthoblast cytonemes in our model. Figure:range_pull The range of xanthoblast cytonemes on $\boldsymbol{X}$. The centre lattice site (marked in purple) represents site $\boldsymbol{X}(r,k)$. The lattice sites labelled in red are those sites $\boldsymbol{X}(i,j)$ that lie in the range of the cytonemes of an X^b occupying site $\boldsymbol{X}(r,k)$ (the sites given by set P). Figure:range_pull_mel The range of xanthoblast cytonemes on $\boldsymbol{M}$. The sites indicated with in orange or yellow are those sites that

can be reached by the range of X^b in $\boldsymbol{X}(i,j)$. The colour of any given site $\boldsymbol{M}(i,j)$ reflects the number of corresponding sites on $\boldsymbol{X}(2i,2j)$, $\boldsymbol{X}(2i-1,2j-1)$, $\boldsymbol{X}(2i-1,2j)$, $\boldsymbol{X}(2i,2j-1)$ that are within range of the cytonemes (marked in red in (**A**)) and thus the probability of the site being chosen in relation to other coloured sites. Specifically, sites labelled in orange are twice as likely to be chosen than those sites marked in yellow. (**C–D**) An example growth event in the vertical direction. Squares represent sites in a given domain. Yellow circles represent cells occupying these sites. (**C**) First a site is chosen at random from each column (marked in red). (**D**) Next a new site is inserted either above or below the first site with probability $\frac{1}{2}$. Each of the chosen sites are marked in dark grey.

To account for the different lattice sizes, each time a growth event in the horizontal (vertical) direction occurs, $\Pi^H_{\boldsymbol{M}}(t)$ ($\Pi^L_{\boldsymbol{M}}(t)$) increases by one and $\Pi^H_{\boldsymbol{X}}(t)$, $\Pi^H_{\boldsymbol{I}}(t)$ ($\Pi^L_{\boldsymbol{X}}(t)$, $\Pi^L_{\boldsymbol{I}}(t)$) increases by two. This means that when a growth event is chosen, the rules described above occur once in the $\boldsymbol{M}$ layer, and twice consecutively in $\boldsymbol{X}$ and $\boldsymbol{I}$. This is to ensure that the domain size remains consistent between the three layers. That is,

$$\Pi^L_{\boldsymbol{X}}(t)\times\Delta_X=\Pi^L_{\boldsymbol{I}}(t)\times\Delta_I=\Pi^L_{\boldsymbol{M}}(t)\times\Delta_M, \tag{59}$$

$$\Pi^H_{\boldsymbol{X}}(t)\times\Delta_X=\Pi^H_{\boldsymbol{I}}(t)\times\Delta_I=\Pi^H_{\boldsymbol{M}}(t)\times\Delta_M, \tag{60}$$

for all time t.

Appendix 2

Mutant implementation

In this text, we describe how the WT fish simulation is altered in our model to replicate changes known to be present in a variety of mutant fish.

Simulating the *shd* mutant

The gene *shady* (*shd*) encodes zebrafish leukocyte tyrosine kinase (Ltk) which plays a role in S-iridophore specification (**Lopes et al., 2008**). As a result, strong *shd* mutants lack S-iridophores. To simulate this defect, we remove all S-iridophores from the initial domain, i.e, we set $N^T_{I^l}(0) = N^T_{I^d}(0) = 0$. Since new S-iridophores are only generated by the proliferation of existing S-iridophores, this means that $N^T_{I^l}(t) = N^T_{I^d}(t) = 0$ for all time t. (Note: As a result the propensities of events 4, 5, 7, 10 and 11, that is, the movement of I^l, movement of I^d, proliferation I^l and I^d, transition of I^d to I^l and transitions of I^l to I^d are always zero during this simulation.)

Simulating the *pfe* mutant

Gene *pfeffer* (*pfe*) (alternatively known as *salz* (*sal*)) encodes for colony stimulating factor one receptor (csf1ra) which plays a role in xanthophore specification and migration. In strong alleles, adult fish exhibit no detectable xanthophores in the body of embryos. To simulate this defect we remove all xanthophores and xanthoblasts from the initial domain, *i.e.* we set $N_{X^b}(0) = N_X(0) = 0$. We also remove the fixed event at stage SP where newly differentiated X cells appear along the horizontal myoseptum. Since new xanthophores are only generated by the proliferation of existing xanthophores, and the differentiation of xanthoblasts this means that $N^T_{X^b}(t) = N^T_X(t) = 0$ for all time t. (Note: As a result the propensities of events 2, 3, 6, 9, 15, that is, the movement of X^b, movement of X, proliferation X and X^b, differentiation of X^b to X and X^b pulling of M are always zero during this simulation.)

Simulating the *nac* mutant

The gene *nacre* (*nac*) encodes transcription factor Mitfa (**Lister et al., 1999**). *nac* mutants lack melanocytes throughout embryonic and larval development (**Lister et al., 1999**). To simulate this defect, we remove all melanocytes from the initial conditions, i.e, we set $N_M(0)$. We also set the propensity of event 8, the differentiation of new M, to be 0 for all time t. Since this is the only way new M can be produced, $N_M(t) = 0$ for all time t. (Note: As a result the propensities of events 1 and 14, that is, the movement of M, and death of M will also always be zero during this simulation.)

Simulating the double mutants: nac;pfe, shd;pfe, shd;nac

To simulate the double mutants we combine the effects of two of the three aforementioned mutation types. For example, to simulate *nac;pfe* we combine the effects of *nac* and *pfe* listed above.

Simulating the *rse* mutant

Rose (*rse*), encodes the Endothelin receptor B1a (**Krauss et al., 2014**) and has been shown to acts cell-autonomously in S-iridophores; homozygous mutants result in a reduction of S-iridophores to approximately 20% of that seen in WT (observed in stage PB and adult fish [**Frohnhöfer et al., 2013**]). To simulate *rse* we change the initial conditions so that $N_{I^d}(0)$ is one fifth of the usual number. In our WT simulations, $N_{I^d}(0)$ is $\Pi^L_{\boldsymbol{I}}(0) \times 3 = 300$ as I^d occupy the

three central rows. In *rse* we set $N_{I^d}(0)$=60, and we place these I^d uniformly at random within the space of the central three rows, since some S-iridophores still appear along the horizontal myoseptum in *rse*. We do this by generating a random number $r \in \mathbb{Z}$ that is uniformly distributed between 24 and 26 (the centre of the horizontal axis is at 25), and a random number k_1 that is uniformly distributed between 0 and $\Pi_{\boldsymbol{I}}^{L}(0) = 100$. If $\boldsymbol{I}(r_1, k_1)$ is empty we place a I^d cell in this site, otherwise we generate a new position until there are 60 cells of type I^d on $\boldsymbol{I}$. Furthermore since new S-iridophores are only produced by proliferation of pre-existing S-iridophores, we also reduce the rate of proliferation of I^d and I^l to one fifth of the usual number. Therefore, we set the propensity of event 7, $\alpha_7(t) = 0.24\frac{N_{I^d}^{T_r}(t)+N_{I^l}^{T_r}(t)}{60\times24}$ for all time t.

Simulating the *sbr* mutant

Adult *sbr* mutants exhibit delayed S-iridophore shape transition changes of dense to loose caused by truncations in Tight Junction Protein 1a (**Fadeev et al., 2015**). To simulate *sbr*, we reduce the rate at which I^d attempts transition to I^l to a fortieth of its usual value. Hence the propensity of event 10 becomes $\alpha_{10}(t) = \frac{N_{I^d}(t)}{60\times24\times40}$. This reduction acts as a proxy for the delay between receiving a signal and changing shape to loose form.

Simulating the *cho* mutant

Mutant larvae with mutation *cho* lack the horizontal myoseptum (**Svetic et al., 2007**). As a result, dense S-iridophores cannot travel through their usual pathway to generate the initial strip of dense S-iridophores at stage PB seen in WT. Instead loose S-iridophores appear later at stage PR, uniformly across the domain. To simulate the effects of *cho*, we remove the initial three rows of I^d so $N_{I^d}(0) = 0$, as S-iridophores cannot appear along the horizontal myoseptum. Furthermore, we remove the fixed event of metamorphic xanthophores appearing along the horizontal myoseptum at stage SP. To simulate the appearance of loose S-iridophores at stage PR, we provide a new fixed event when the simulation reaches stage PR. At this point, we place 300 (the same number that usually initially occupies the domain in WT) loose S-iridophores in sites uniformly at random across $\boldsymbol{I}$ at stage PR. We do this by generating a random number r between 0 and $\Pi_{\boldsymbol{I}}^{L}(t_{PR})$ and a random number k between 0 and $\Pi_{\boldsymbol{I}}^{H}(t_{PR})$, where t_{PR} is the time when the simulation first enters stage PR. If $\boldsymbol{I}(r, k)$ is empty we place a I^l cell in this site, otherwise we generate a new r, k and continue until there are 200 cells of type I^l on $\boldsymbol{I}$.

Simulating the *seurat* mutant

Homozygous *seurat* mutants develop fewer adult melanocytes, thus forming irregular spots rather than stripes. This phenotype arises from lesions in the gene encoding Immunoglobulin superfamily member 11 (Igsf11) (**Eom et al., 2012**) which encodes a cell surface receptor containing two immunoglobulin-like domains which is expressed autonomously by the melanocyte lineage. Igsfl1 promotes the migration and survival of these cells during adult stripe development as well as mediating adhesive interactions in vitro.

To model *seurat* we reduced the rate at which melanocytes could differentiate to a twentieth of the usual rate. Hence the propensity of event eight becomes $\alpha_8(t) = 0.05 \times \frac{N_{I^d}(t)}{2\times60\times24}$. This was to reflect the inhibition of migration of melanoblasts across the domain and increased the rate of attempted melanocyte death to one hundred times per day (usually once per day). Hence the propensity of event 12 becomes $\alpha_{12}(t) = \times\frac{N_M(t)}{100\times60\times24}$. No other interactions were altered.

Simulating the *leo* mutant

The gene *leo* encodes Connexin39.4 (Cx39.4). The *leo* mutant displays a spotted pattern across the flank of the fish. In Section "An *in silico* investigation into the function of the *leo* gene" of the main text we describe how we derive the following hypotheses about the impacts of a mutation in the *leo* gene;

- Hypothesis 1: Melanocytes are not repelled by xanthophores
- Hypothesis 2: Xanthophores do not promote the survival of melanocytes in the long range.
- Hypothesis 3: Xanthophores do not promote the change of S-iridophores from dense to loose in the long range.
- Hypothesis 4: Melanocytes lose death signals from local dense S-iridophores and as a result, can differentiate in dense S-iridophore zones.
- Hypothesis 5: Melanocytes lose directed signalling from S-iridophores and hence in the absense of xanthophores differentiate randomly.

To model hypothesis 1, we change the parameter r_{mx}, the parameter governing repulsion of melanocytes from xanthophores to 0. To incorporate hypothesis two we remove the criteria for successful melanocyte death given in ***Equation 50*** in Appendix 1. To model hypothesis 3, we reduce the rate of successful signalling of xanthophores to iridophores to change to loose form. To do this, if a S-iridophore transition is chosen to occur then a number distributed uniformly at random is generated. If this number is less than 0.5, then normal transition signalling occurs (as described in Appendix 1). If the number is greater than 0.5 then the xanthophores send a signal for S-iridophores to transition to dense, that is, the number N_I, X^S is changed to five and the number N_I, X^L is changed to 2. To model hypotheses 4 and 5, we change the melanocyte differentiation success as follows. A melanocyte successfully differentiates into a position on the lattice, if there are no xanthophores on the domain, or if there are xanthophores on the domain and there are three times as many xanthophores in the long range than the number of melanocytes in the long-range.

Appendix 3

Quantitative measures

In this text, we describe in more detail how we take quantitative measurements of our simulations.

Tortuosity

To determine the tortuosity of the X0 stripe, for a specific time point in our simulations the following steps are taken.

- An occupancy matrix $\hat{\boldsymbol{X}}$ of zeros and ones is generated such that a matrix entry $\hat{\boldsymbol{X}}(i,j) = 1$ if $\boldsymbol{X}(i,j) = X$ and 0 otherwise. Sites (i,j) such that $\hat{\boldsymbol{X}}(i,j) = 1$ are yellow and are white otherwise.
- Representative matrix $\hat{\boldsymbol{X}}$ is 'cleaned' using matlab functions bwmorph($\hat{\boldsymbol{X}}$,'clean') and bwmorph($\hat{\boldsymbol{X}}$,'bridge') consecutively. 'Clean' removes any anomalous xanthophores that are not connected (adjacent to) other xanthophores. 'Bridge' adds extra xanthophores where there are holes in the population pattern.
- An algorithm is applied to $\hat{\boldsymbol{X}}$ to create the outline of the X0 stripe. The algorithm to create the top line (l_t) is given below in Algorithm 1. A similar algorithm is used to generate the bottom line (l_b).
- A line L that represents the middle of the stripe is given by $l_m(i) = \frac{l_t(i)+l_b(i)}{2}$ for $i = 1, ...\Pi_{\boldsymbol{X}}^{L}$.
- We smooth L for analysis by applying matlab function smooth. 'Smooth' smooths the data in the column vector y using a moving average filter.
- Finally, we calculate the tortuosity of the line by computing the total length of L divided by the distance (algorithm given below in Algorithm 2).

Algorithm 1. Algorithm to generate a representative line for the top of stripe X0 (l_t)

$j = \left\lceil \frac{\Pi_{\boldsymbol{X}}^{H}}{2} \right\rceil$ ▷*Initialise X0 interstripe search in the center.*
for $i = 1 : \Pi_{\boldsymbol{X}}^{L}$
  Strike = 0 ▷*Initialise the number of consecutive empty sites in a column to be zero.*
  if $\hat{\boldsymbol{X}}(i,j) \neq 0$ or ($\hat{\boldsymbol{X}}(i,j) = 0$ and $\hat{\boldsymbol{X}}(i,j+1) \neq 0$) then
  while $strike$<2 and j<$\Pi_{\boldsymbol{X}}^{H}$ do ▷*If we are near to other X, keep moving up to determine the upper bound of the interstripe.*
  $j = j + 1$
  if $\hat{\boldsymbol{X}}(i,j) = 0$ then
  $strike = strike + 1$ ▷*Increment the number of consecutive empty sites in a column by one.*
  else
  $strike = 0$
  end if
  end while
  $j = j - 2$ *Remove the two consecutively empty sites from the total count.*
  else
  while$\hat{\boldsymbol{X}}(i,j) = 0$ and j>$\Pi_{\boldsymbol{X}}^{H}$ do ▷*If we have overshot the interstripe keep moving downwards until we reach the top of the interstripe.*
  $j = j - 1$
  end while
  end if
  $l_t(i) = j$
end

Algorithm 2. Algorithm to calculate the tortuosity ($tort$) of line L generated from a simulated X0

total = 0
for $i = 2 : \Pi_{\boldsymbol{X}}^{L}$
  $total = total + \sqrt{(l_m(i) - l_m(i-1))^2 + 1}$ ▷*Compute the length of L using the euclidean distance between consecutive points.*
end
$tort = \frac{total}{\Pi_{\hat{\boldsymbol{X}}}^{L}}$

To measure the tortuosity of the stripes of real fish, a photograph was taken of the fish at stage J+. Next, matlab function 'getpts' was used to mark the outline of the X0 stripe by a set of points $[\underline{x},\underline{y}]$ where vector $\underline{x}$ is length k. The tortuosity of this line was then computed using Algorithm 3.

Algorithm 3. Algorithm to calculate the tortuosity ($tort$) of line L generated from a real fish

$total = 0$
for $i = 2 : k$
  $total = total + \sqrt{(x(i)-x(i-1))^2+(y(i)-y(i-1))^2}$
end
$tort = \frac{total}{\sqrt{(x(1)-x(k))^2+(y(i)-y(k))^2}}$

▷ *Compute the length of L using the euclidean distance between consecutive points.*

X0 interstripe width

To determine the width of the X0 interstripe, used in Section "Quantitative analysis of simulations" of the main text, first, occupancy matrices of $\hat{\boldsymbol{X}}$ and $\hat{\boldsymbol{I}}$ of zeros and ones is generated such that a matrix entry $\hat{\boldsymbol{X}}(i,j)=1$ if $\boldsymbol{X}(i,j)=X$ and 0 otherwise. Similarly $\hat{\boldsymbol{I}}(i,j)=1$ if $\boldsymbol{I}(i,j)=I^d$ and 0 otherwise. We then generate l^t and l^b using either $\hat{X}$ or $\hat{I}$ as described in Section "Necessity of S-iridophore assumptions". From these values, we generate IS (X0 interstripe width) by computing;

$$IS=\sum_{i=1}^{\Pi_X^L}\frac{l_t(i)-l_b(i)}{\Pi_X^L}. \tag{61}$$

Note that the X0 interstripe width is computed using whichever is more appropriate of X on $\boldsymbol{X}$ and I^d on $\boldsymbol{I}$ given the mutation. For example, *pfe* does not have xanthophores, so we would use the distribution of I^d on $\boldsymbol{I}$ to determine the width of X0 in this case.

Adapting the pair correlation function (PCF)

PCFs characterise spatial patterns by calculating a numerical value for the deviation from the situation in which the same number of agents are distributed uniformly at random. In this paper, we use the square uniform PCF (***Gavagnin et al., 2018***) developed for on-lattice systems of agents where distance is measured using the uniform norm. This PCF was originally developed for determining pair correlation between single agents types on a lattice, however, in a technique similar to ***Dini et al., 2018*** we adapt it here so that it can also be used for identifying correlation between two different types of agents (cell types). First we define the PCF. For each distance m the PCF at distance m is given by:

$$\mathrm{PCF}(m)=\frac{c^{C_1,C_2}(m)}{\mathbb{E}[\hat{c}^{C_1C_2}(m)]} \tag{62}$$

where $c^{C_1,C_2}(m)$ is defined as the number of cells of type C_1 we would expect to find at distance m from cells of type C_2 using the uniform metric under zero flux boundary conditions. $\mathbb{E}[\hat{c}^{C_1,C_2}(m)]$ if the cells were positioned uniformly at random on the domain. This can be calculated by counting the number of agents of this distance manually from the lattice. To compare cell type M with cells that lie on domains other than $\boldsymbol{M}$ we must transform $\boldsymbol{M}$ into a matrix $\hat{\boldsymbol{M}}$ of size $\Pi_X^H\times\Pi_X^L$ where $\hat{\boldsymbol{M}}(r,k)=M$ if $\boldsymbol{M}(\lceil\frac{r}{2}\rceil,\lceil\frac{k}{2}\rceil)=M$ and 0 otherwise. Hence the set of site positions distance m from each other on any domain $D\in\{\boldsymbol{X},\hat{\boldsymbol{M}},\boldsymbol{I}\}$ under zero flux boundary conditions can be given by

$$S_m=\{(r,k),(i,j)\in(\Pi_X^L,\Pi_X^H),(\Pi_X^L,\Pi_X^H)\,|\,\max\{|r-i|,|k-j|\}=m\}. \tag{63}$$

Therefore

$$c^{C_1,C_2}(m)=\begin{cases}\sum_{(i,j),(r,k)\in S_m}1_{D_1(r,k)=D_2(i,j)=C}, \text{where} C_1=C_2=C,\\ \sum_{(i,j),(r,k)\in S_m}1_{D_1(r,k)=C_1 \text{and} D_2(i,j)=C_2}+1_{D_2(r,k)=C_2 \text{and} D_1(i,j)=C_1} \text{otherwise},\end{cases} \tag{64}$$

where C_1 lies on D_1, C_2 lies on D_2 and $D_1,D_2\in\{\boldsymbol{X},\hat{\boldsymbol{M}},\boldsymbol{I}\}$. Note that $|S_m|$ is the number of site pairs distance m from one another. $|S_m|$ is computed in ***Gavagnin et al., 2018*** and is given by

$$|S_m|=4m\Pi_{\boldsymbol{X}}^H\Pi_{\boldsymbol{X}}^L-3(\Pi_{\boldsymbol{X}}^H+\Pi_{\boldsymbol{X}}^L)m^2+2m^3. \tag{65}$$

To determine $\mathbb{E}[\hat{c}^{C_1,C_2}(m)]$ there are three cases.

Case 1: $C_1=C_2=C\in\{M,X,I^d,I^l,X^b\}$ lies on domain $D\in\{\boldsymbol{M},\boldsymbol{X},\boldsymbol{I}\}$ (with volume exclusion on D). For example spatial correlation of M with respect to other M. This is the case discussed in ***Gavagnin et al., 2018***. In this case

$$\mathbb{E}[\hat{c}^{C,C}(m)]=\mathbb{P}(\text{Two agents of type } C \text{ are chosen from domain } D)|S_m|. \tag{66}$$

$$\mathbb{P}(\text{Two cells of type } C \text{ are chosen from domain } D)=\frac{N_C^T(N_C^T-1)}{(\Pi_{\boldsymbol{X}}^H\Pi_{\boldsymbol{X}}^L)(\Pi_{\boldsymbol{X}}^H\Pi_{\boldsymbol{X}}^L-1)}. \tag{67}$$

This is because there are N_C^T ways to choose a cell of type C and N_C^T-1 ways to choose a cell of type C given one has been chosen already. There are $\Pi_{\boldsymbol{X}}^H\Pi_{\boldsymbol{X}}^L$ possible positions for the first cell and then $\Pi_{\boldsymbol{X}}^H\Pi_{\boldsymbol{X}}^L-1$ possible positions for the second cell due to volume exclusion on domain D.

Case 2: $C_1\neq C_2$, where $C_1,C_2\in\{X,I^d,I^l,X^b\}$ both lie on domain $D\in\{\boldsymbol{X},\boldsymbol{I}\}$ (with volume exclusion on D). For example spatial correlation of X with X^b. In this case

$$\mathbb{E}[\hat{c}_d^{C_1,C_2}(m)]=\mathbb{P}(\text{ One cell of type } C_1 \text{ and one cell of type } C_2 \text{ are chosen from domain } D)2|S_m|, \tag{68}$$

Notice, $|S_m|$ is multiplied by two here because for each pair $(i,j),(r,k)\in S_m$ there are two positions C_1 and C_2 can take and be distance m away from each other. Specifically these are the cases $D_1(i,j)=C_1$, $D_2(r,k)=C_2$ and $D_2(i,j)=C_2$, $D_1(r,k)=C_1$. We can compute

$$\mathbb{P}(\text{ One cell of type } C_1 \text{ and one cell of type } C_2 \text{are chosen on D})=\frac{\mathrm{N}_{\mathrm{C}_1}^{\mathrm{T}}\mathrm{N}_{\mathrm{C}_2}^{\mathrm{T}}}{(\Pi_{\boldsymbol{X}}^{\mathrm{H}}\Pi_{\boldsymbol{X}}^{\mathrm{L}})(\Pi_{\boldsymbol{X}}^{\mathrm{H}}\Pi_{\boldsymbol{X}}^{\mathrm{L}}-1)}. \tag{69}$$

This is because there are $N_{C_1}^T$ ways to choose a cell of type C_1 and $N_{C_2}^T$ ways to choose a cell of type C_2. There are $\Pi_{\boldsymbol{X}}^H\Pi_{\boldsymbol{X}}^L$ possible positions for the first cell and then $\Pi_{\boldsymbol{X}}^H\Pi_{\boldsymbol{X}}^L-1$ possible positions for the second cell due to volume exclusion on domain D.

Case 3: C_1, C_2 where $C_1,C_2\in\{M,X,I^d,I^l,X^b\}$ lie on domains D_1, $D_2\in\{\hat{\boldsymbol{M}},\boldsymbol{X},\boldsymbol{I}\}$ where $D_1\neq D_2$, for example the spatial correlation of X and I^d. In this case

$$\mathbb{E}[\hat{c}^{C_1,C_2}(m)]= \mathbb{P}(\text{ One cell of type } C_1 \text{ is chosen on } D_1 \text{ and one cell of type } C_2 \text{ is chosen on } D_2)2|S_m|, \tag{70}$$

where

$$\mathbb{P}(\text{ One cell of type } C_1 \text{ is chosen on } D_1 \text{ and one cell of type } C_2 \text{ is chosen on } D_2)=\frac{N_{C_1}^TN_{C_2}^T}{(\Pi_{\boldsymbol{X}}^H\Pi_{\boldsymbol{X}}^L)^2}. \tag{71}$$

This is because there are $N_{C_1}^T$ ways to choose a cell of type C_1 and $N_{C_2}^T$ ways to choose a cell of type C_2. There are $\Pi_{\boldsymbol{X}}^H\Pi_{\boldsymbol{X}}^L$ possible positions for the first cell of type C_1 and $\Pi_{\boldsymbol{X}}^H\Pi_{\boldsymbol{X}}^L$ possible positions for the second cell of type C_2 since they lie on different domains (no volume exclusion). A summary of all these cases can be given as follows;

- $C_1=C_2=C$ lies on domain D (with volume exclusion on D). For example, spatial correlation of M with respect to other M.

$$PCF(m) = \frac{c^{C,C}(m)(\Pi_X^H \Pi_X^L)(\Pi_X^H \Pi_X^L - 1)}{(4m\Pi_X^H \Pi_X^L - 3(\Pi_X^H + \Pi_X^L)m^2 + 2m^3)(N_C^T(N_C^T - 1))} \tag{72}$$

- $C_1 \neq C_2$, C_1, C_2 lie on domain D (with volume exclusion on D). For example spatial correlation of X with respect to X^b.

$$PCF(m) = \frac{c^{C_1,C_2}(m)(\Pi_X^H \Pi_X^L)(\Pi_X^H \Pi_X^L - 1)}{2(4m\Pi_X^H \Pi_X^L - 3(\Pi_X^H + \Pi_X^L)m^2 + 2m^3)(N_{C_1}^T(N_{C_2}^T))} \tag{73}$$

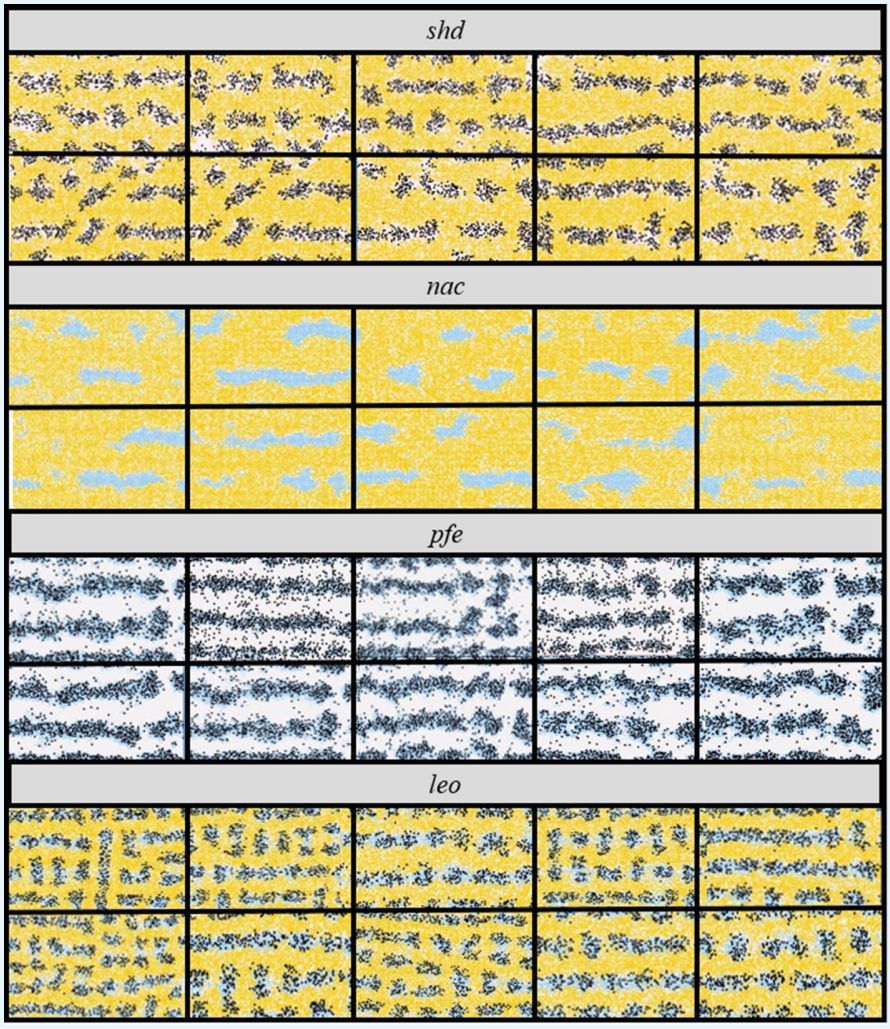
- $C_1 \neq C_2$, C_1, C_2 lie on domains D_1, D_2 where $D_1 \neq D_2$ for example spatial correlation of X with respect to M.

$$PCF(m) = \frac{c^{C_1,C_2}(m)(\Pi_X^H \Pi_X^L)^2}{2(4m\Pi_X^H \Pi_X^L - 3(\Pi_X^H + \Pi_X^L)m^2 + 2m^3)(N_{C_1}^T(N_{C_2}^T))} \tag{74}$$

Appendix 4

Testing robustness

Due to the abundance of parameters and cell–cell interactions necessary to capture what is known biologically about zebrafish pigment pattern formation, it is not feasible to perform an exhaustive parameter sweep to demonstrate the robustness of the model. Instead, as a test of robustness, we perform a rigorous robustness analysis by carrying out one hundred repeats of all WT and mutant simulations with perturbed parameter values chosen uniformly at random from the range 0.75–1.25 of their described value. The value of each parameter is sampled uniformly from this region, independentally for each parameter and each repeat. Ten of these randomly sampled repeats for *shd, nac, pfe* and *leo* are given in ***Appendix 4—figure 1***. We observe, for all one hundred repeats, that small perturbations to the rates still generate consistent patterning, demonstrating the robustness of the model.



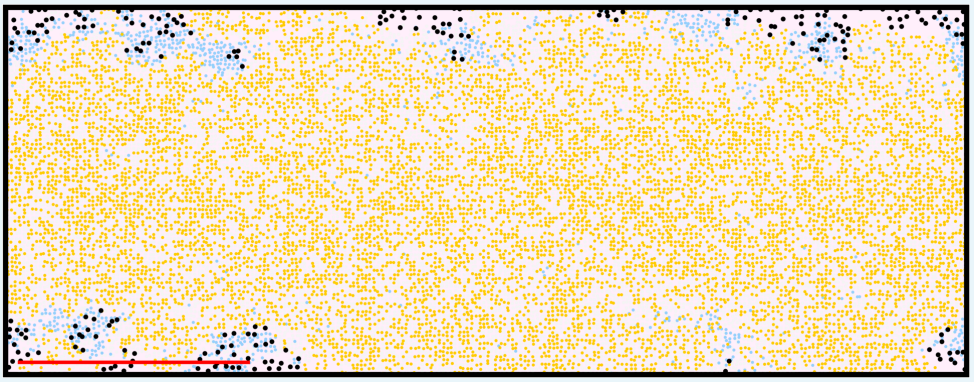
Appendix 4—figure 1. Example *shd, nac, pfe* and *leo* simulations at stage J+ when the parameters governing the rate of proliferation, movement, differentiation and death are perturbed. Each rectangle is an example simulation at stage J+ for which each rate parameter is perturbed to $1+x$ times its normal value. The value x is chosen uniformly at random from the interval $[-0.25, 0.25]$.

Appendix 5

Predicting pattern formation for an *seurat/sbr*cross

In this section, we give an example of how the model can be manipulated to investigate aspects of pattern development and to predict the outcome of mutant pigment patterns.

To the best of our knowledge, the adult pattern of crossed *seurat* and *sbr* mutants have not have been published previously. By changing the parameters so that we match both *sbr* and *seurat* we can predict pattern formation for a double mutant *seurat/sbr*. As previously stated (in more detail) in Appendix 2, the mutation in *seurat* affects melanocyte differentiation (***Eom et al., 2012***) and the mutation in *sbr* affects the dense-to-loose S-iridophore transition (***Fadeev et al., 2015***). To simulate the cross, we simply incorporate the effects of both mutations simultaneously. The results at stage J+ is given in ***Appendix 5—figure 1***. Our model predicts that when both of these mutations occur, dense S-iridophores and associated xanthophores would cover most of the flank of the fish. A few melanocytes associated with loose S-iridophores survive at the very dorsal and ventral regions of the fish.



Appendix 5—figure 1. Prediction for the pattern of a *seurat/sbr* cross at stage J+.

6.3. Conclusions

The aim of this paper is to investigate zebrafish stripe formation using a comprehensive mathematical model. By adopting a bottom-up individual-based modelling approach we are able to capture all of the key aspects of the literature in our model and to determine the sufficiency of the interactions for generating the patterns.

The most significant findings of this paper are the predictions made regarding pattern formation. These findings will guide future experimentation and aid the understanding of zebrafish stripe formation as well as pattern formation more generally. In particular we predict that the initial interstripe is not the only factor determining stripe orientation, nor the most important factor. This is explored in more detail in Chapter 8 in which we investigate the effects of growth during development.

We also make predictions regarding functionality of the *leo* gene. This work provides testable hypotheses for future investigation. For example, recently, Usui *et al* published a study into function of connexins 39.4 (*leo*) and connexin 41.8 (*luc*) [71]. They generated transgenic lines expressing *leo* and *luc* in melanophores and or xanthophores individually which they subsequently used to identify the minimal requirements of connexin expression in pigment cells for stripe formation.

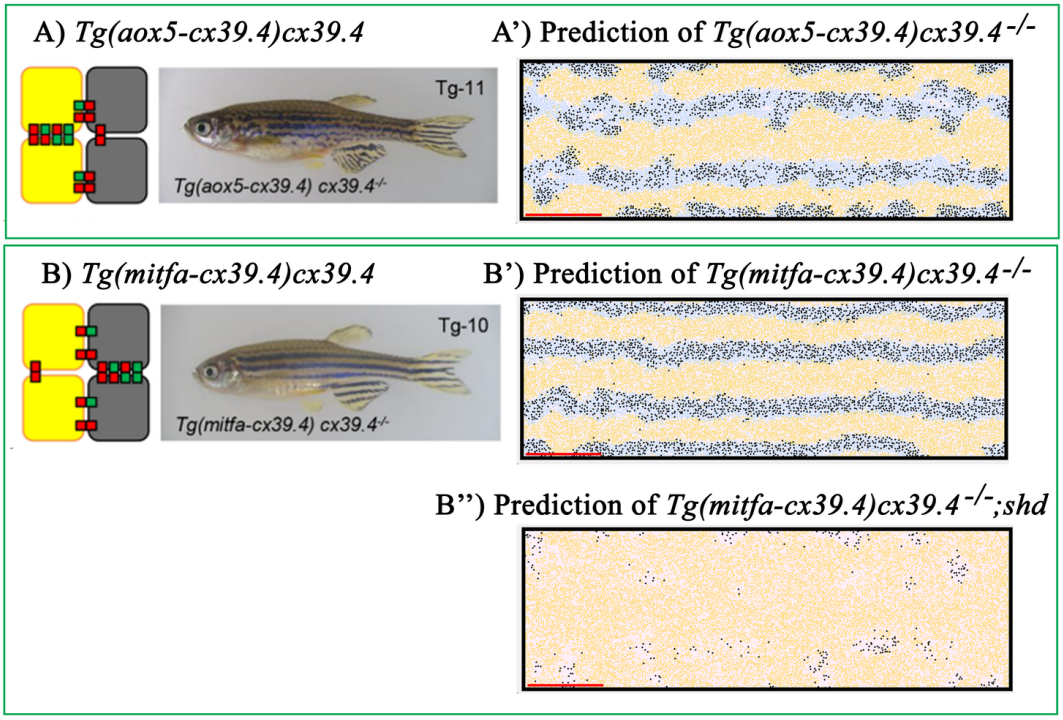


Figure 6-1: ***Using our model to make predictions about transgenic leo mutants.*** *(A),(B) Transgenic fish that can only produce the leo gene (Connexin 39.4) in (A) xanthophores (Tg(aox5-cx39.4)cx39.4$^{-/-}$), (B) melanophores (Tg(mitfa-cx39.4)cx39.4$^{-/-}$) as shown by Usui et al [71]. (A'),(B') Predictions of the pigment pattern at stage J+ using our mathematical model with our predictions of the functionality of the leo gene given in the paper which forms the basis of this Chapter for transgenic fish: (A') Tg(aox5-cx39.4)cx39.4$^{-/-}$, (B) Tg(mitfa-cx39.4)cx39.4$^{-/-}$. (B") A testable prediction of the patterns of double mutant: Tg(mitfa-cx39.4)cx39.4$^{-/-}$;shd*

We can use their findings in conjunction with our model in order to test our predictions about the function of *leo* theoretically. Fig. 6-1(A) displays the adult pattern of transgenic fish *Tg(aox5-cx39.4)cx39.4*$^{-/-}$. The genetics of this fish means that it

Genotype	**Actual Mel/mm^2**	**Predicted Mel/mm^2**
WT	62.0	75.8
Tg(aox5-cx39.4)cx39.4$^{-/-}$	43.3	43.8
Tg(mitfa-cx39.4)cx39.4$^{-/-}$	52.3	49.4
cx39.4$^{-/-}$ (leo)	36.5	33.9

Table 6.1: *A comparison between the actual numbers of melanophores per mm^2 observed on WT, Tg(mitfa-cx39.4)cx39.4$^{-/-}$, Tg(mitfa-cx39.4)cx39.4$^{-/-}$ and leo as determined by Usui et al [71] and the mean predicted melanophores per mm^2 of our simulations, averaged over 50 repeats. Melanophores per mm^2 in our simulation were computed by calculating the number of melanophores in a mm^2 subsection of the stripes. Since each lattice site represents 40μm of the real fish, this corresponds to an area of 25 × 25 lattice sites.*

can only produce the *leo* gene in xanthophores. The adult fish displays undulating stripes with a reduced number of melanophores per mm^2 compared to WT (Table 6.1). We can generate the same 'genetics' in our mathematical model by only implementing the predicted *leo* rules that include melanocytes. The section: 'An *in silico* investigation into the function of the *leo* gene' in the paper provided within this Chapter details these hypotheses. The set of hypotheses that hold true in the genetics of *Tg(aox5-cx39.4)cx39.4$^{-/-}$* comprise hypotheses 1, 2, 4 and 5. We show in Fig. 6-1(A') a representative simulation of the fish at J+ with the rules as described. The pattern that is generated is also an undulating striped pattern. We also see the same comparative decrease in the number of melanophores per mm^2 (Table 6.1). Fig. 6-1(B) displays the the adult pattern of transgenic fish *Tg(mitfa-cx39.4)cx39.4$^{-/-}$*. This fish only produces *leo* gene in melanophores and displays normal striping, albeit with a slightly reduced number of melanophores compared to WT (Table 6.1). We can generate the same 'genetics' in our mathematical model by only implementing the predicted *leo* rules that include xanthophores. Similarly, the genetics of *Tg(mitfa-cx39.4)cx39.4$^{-/-}$* comprise the set of hypotheses 1 and 2. We show in Fig. 6-1(B') that in doing so we also produce a striping pattern as seen in 6-1(B). Similarly we see the same comparative decrease in the mean number of melanophores per mm^2 (Table 6.1). These results further suggest that our predictions for *leo* are valid.

Satisfied that our predictions hold, we can also use our model to guide future experimentation, by suggesting analysis that can test our predictions experimentally. In Fig. 6-1(B') we give a representative simulation of *Tg(mitfa-cx39.4)cx39.4$^{-/-}$* without iridophores. This is equivalent to *Tg(mitfa-cx39.4)cx39.4$^{-/-}$;shd*. We predict that despite the transgenic fish looking normal, that if our rules regarding *leo* are correct then *Tg(mitfa-cx39.4)cx39.4$^{-/-}$;shd* should display a reduction in melanophores with no obvious spots. This is contrast to *shd* which displays melanophore spots in a sea of xanthophores.

This study gives just one example of how our mathematical model can aid future investigations. In the next chapter we will generate a new method of repeatedly imaging zebrafish during the pattern development stage which we can use to parameterise the frequencies of some of the events that occur during pattern development, such as the rate of cell proliferation, death and interactions of iridophores. In doing so we can use our model iteratively, generating real experimental data to improve the model and subsequently using the results of the model to fuel future experimentation.

Chapter 7

A new anaesthetic protocol for metamorphic zebrafish

7.1. Overview

In Chapter 6 we showed that we were able to build a comprehensive model of zebrafish stripe formation. Whilst most of the parameters for the model are known, some remain unexplored. In order to parameterise our model further we are required to repeatedly image zebrafish during the pattern development period. In this way we are able to track the frequency of actions such as cell movement, cell death for example over a period of days which we can then use to parameterise our model accordingly.

This chapter is dedicated to determining a suitable anaesthetic protocol for anaesthetising and recovering zebrafish during juvenile stages during pattern metamorphosis (8 mm SL - 16 mm SL). The protocol determined appropriate in this chapter is used in Chapter 8 to track the relative position of the X0 interstripe along the dorso-ventral axis during pattern metamorphosis as part of an in-depth study of the impact of growth. This chapter comprises the paper written by myself (Jennifer Owen) and Robert Kelsh, and was accepted for publication to PLOS ONE in December 2020.

In the paper we identify the juvenile stage during zebrafish development as a period in which fish become sensitive to the anaesthetic MS-222: we show that despite MS-222 being the most commonly used fish anaesthetic [107, 108, 109], that no standard concentration of MS-222 can provide reliable anaesthesia for a prolonged test period (30 minutes) during these stages. We next test six published anaesthesia protocols for a shorter period of 10 minutes and assess reliable anaesthesia on juvenile zebrafish subject to a single dose. We show that four of the six protocols are suitable; high recovery rate and no movement under anaesthetic, allowing us to eliminate the remaining two. We test the remaining four anaesthetic methods for repeated trials every 4 days against pre-defined success measures. Our results indicate that one protocol in particular: 2-phenoxythenol at a concentration of 0.085%, where fish are maintained anaesthetised on a cotton pad doused in the anaesthetic, is the most suitable method for use in repeated anaesthesia during these stages. For example, the fish would lose touch responsivity quickly, recover quickly and did not move under anaesthetic. Repeating the trial using 2-phenoxyethanol every 2 days we determine that this method is also suitable for juvenile stages for frequent repeated anaesthesia out of all of the methods trialled.

Our results provide a significant refinement to the method used to anaesthetise juvenile zebrafish during development, thus helping minimise suffering, in one of the most scientifically important vertebrate model systems.

7.2. Outline of the paper

Section I provides an outline of the need for a protocol to anaesthesise juvenile fish as well as giving a brief review of previous studies regarding the anaesthesis of zebrafish more generally. In Section II of the paper we outline the Materials and Methods. In Section III we describe the results of our experiments. Following a series of experiments,

we determine that using 2-phenoxyethanol at a concentration of 0.085% wherein fish are maintained under anaesthetic on a cotton pad to repeatedly anaesthetise fish is the most suitable method. We discuss the findings in Section IV and finish with conclusions in Section V. The supplementary material contains tables outlining the details of all multi-linear regression analysis performed as well as some additional plots.

Appendix B: Statement of Authorship

This declaration concerns the article entitled:			
A suitable anaesthetic protocol for metamorphic zebrafish			
Publication status (tick one) **Draft manuscript** [ ] **Submitted** [ ] **In review** [ ] **Accepted** [ ] **Published** [X]			
Publication details (reference)	Journal: PLOS ONE, 052021 Authors: Jennifer P. Owen and Robert N. Kelsh		
Copyright status (tick the appropriate statement) I hold the copyright for this material [ ] Copyright is retained by the publisher, but I have been given permission to replicate the material here [X]			
Candidate's contribution to the paper (provide details, and also indicate as a percentage)	All the experiments and statistical calculations have been performed by the author of the thesis (100%). All authors contributed equally to the presentation of the content (50%).		
Statement from Candidate	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature.		
Signed	Jennifer Owen	**Date**	07/06/2021

A suitable anaesthetic protocol for metamorphic zebrafish

Jennifer Owen[1] , Robert N Kelsh[1]*
1 Department of Biology and Biochemistry, University of Bath, Bath, United Kingdom
*Corresponding author e-mail: bssrnk@bath.ac.uk

7.3. Abstract

Zebrafish are frequently used as a means to investigate development. These studies increasingly require repeated anaesthesia of zebrafish during juvenile (i.e. metamorphic) stages. The effects of anaesthesia during this time remain poorly studied. The aim of this study was to develop a reliable method that can be used for frequently repeated anaesthesia during juvenile stages. Initially, we assessed different concentrations of MS-222, the most commonly used fish anaesthetic, for 30 minute anaesthesia with recovery. We showed that suitable MS-222 doses could be identified for the smallest (7mm) and largest (20mm) fish. However, we found that juvenile fish within a specific metamorphic window (sized between 8 - 16 mm) were vulnerable to MS-222 and no standard concentration of MS-222 provided reliable anaesthesia under these conditions. Hence we focussed our efforts on identifying a protocol for these stages. We tested six different published anaesthesia protocols P1 - P6 where P1, P2 corresponds to 0.01% MS-222, P3, P4: 0.085% 2-phenoxyethanol and P5, P6: 0.00025%/ 0.0050% Propofol/Lidocaine. In protocols P1, P3, P5 fish were maintained by immersion, whilst in P2, P4 and P6: fish were maintained on an anaesthetic-doused cotton-pad. We assessed reliable anaesthesia using 10 fish for 10 minutes, with full recovery. Our data allowed us to eliminate two of these protocols as unsuitable for short term anaesthesia with recovery of juvenile fish. Extending these studies to explore repeated anaesthesia at 4 day intervals for 20 days under the remaining four protocols, we showed that P1 and P4 were both suitable for repeated anaesthesia, and that P4 was most suitable for imaging. We confirmed that P4 remained suitable when the frequency of anaesthesia was increased to every 2 days. We conclude that this protocol provides a refinement to the current protocol for repeated anaesthesia with recovery of juvenile zebrafish in the vulnerable metamorphic window. .

7.4. Introduction

Laboratory zebrafish (*Danio rerio*) were originally developed as a model organism for developmental biology, but their use has since spread to incorporate congenital and degenerative diseases, regeneration, and toxicology [110]. For example, within developmental biology, zebrafish have become a key model organism for the understanding of pigment pattern formation [27, 61]. Zebrafish generate striking horizontal blue and yellow stripes during the metamorphic period between 21-70 days post fertilisation (dpf) which is caused by the self-organisation of pigment producing cell types across the body of the skin [46]. Many investigations have been undertaken in the last decade in order to understand the cell and molecular biology underpinning the self-organisation of these cells during this time [18, 33, 46, 72].

Central to documenting the development of the pigment pattern is repeated photography of the pigment cell distribution in juvenile fish throughout this period; these images can then be used to track the appearance, migration and death of cells on an individual zebrafish during this period. Photography requires general anaesthesia of the fish, and this needs to be repeatedly applied to document the progression of pattern

formation. Other general non-invasive techniques that may require repeated anaesthesia over a short period of time include imaging of developmental processes [111], mucus collection and other non-lethal sampling methods [112, 113].

General anaesthesia is defined as a temporary loss of sensation and awareness through depression of the central nervous system. This state may be followed by different levels of analgesia (absence of pain) and muscle relaxation. The use of an appropriate anaesthetic protocol for scientific procedures is important. Fish that are not anaesthetised properly may experience stress and pain during the procedure compromising animal welfare. Poor anaesthesia can also lead to data variability; for example, fish that move during imaging can impede imaging. On the other hand, fish that are anaesthetised too deeply may not recover, cutting time-series studies short.

To date, a limited number of anaesthetics have been rigorously tested, with almost all work focusing on the very young (embryonic/early larval, <5 dpf [114]) and adult (>100 dpf) [115, 116, 117, 118, 107] fish, with juvenile stages being neglected. MS-222 (tricaine methanesulfonate) is the anaesthetic that has been the most largely used by the scientific community [119]. MS-222 is a muscle relaxant that blocks sodium and to a lesser degree potassium currents in nerve membranes [120]. It is a water-soluble anaesthetic commonly used for fishes and other cold-blooded animals and is considered safe [108, 107, 109]. However, recent studies have suggested that MS-222 has limitations. Without the correct dose and exposure time, MS-222 in adult zebrafish can have adverse side effects such as aversion and stress induction [118], epidermal and corneal lesions, hypoxemia, decreased heart rate and higher mortality under long-term sedation [121]. Such effects can be an issue when deeper stages of anaesthesia and long duration procedures are needed, as well as when repeated imaging is required [114].

Aside from MS-222, a limited number of other anaesthetics have also been used or tested on adult zebrafish. These include; 2-phenoxyethanol [116], etomidate [122], propofol [122, 123], lidocaine [122, 123] and ketamine [122], although so far there is no conclusive evidence to suggest that these compounds are better than MS-222 itself. Other alternatives consist of combinations of these compounds. For example, in comparison with using MS-222 alone, low doses of isoflurane with MS-222 have been shown to prolong the length of safe anaesthesia (anaesthesia duration without death) and to increase the speed of recovery in adult zebrafish, with minimal effects on the heart rate [121]. More recently, Valentim et al, trialled various combinations of analgesics with anaesthetics, of which they found that propofol with lidocaine was the most successful combination in anaesthetising and recovering adult fish [122, 123]. They found that there were various advantages and disadvantages when comparing the combination with MS-222 alone and advocated tailoring anaesthetic regime to experimental requirements [123].

In summary, whilst there has been an increase in the use of juvenile zebrafish in developmental studies, there is no rigorously tested protocol for anaesthesia in juvenile stages. Furthermore, whilst MS-222 has been identified as a reasonable anaesthetising agent in adult fish [115, 119, 107], its adverse effects leave uncertainty as to whether it is the most refined option for juvenile fish. In the light of new alternatives, it is possible that other treatments are more suitable at this sensitive developmental stage, especially in the context of repeated treatment.

In this three-part study, we determine an anaesthetic protocol for repeatedly anaesthetising and recovering juvenile fish. Firstly, we investigate the effects of different concentrations of MS-222 as suitable for single-use anaesthesia with recovery for juvenile fish. From this investigation, we demonstrate a juvenile developmental time window within which zebrafish are particularly sensitive to MS-222. Because fish of in-

termediate length were highly sensitive to the concentration of MS222, it is difficult to define appropriate standardised procedures for anaesthesia during this developmental stage. Secondly, we focus on this sensitive time window, using a shorter anaesthesia window (10 mins), broadening our range of anaesthetics, and trialling different methods for maintenance of anaesthesia, to determine an effective anaesthesia protocol for permitting a workable duration of anaesthesia with good recovery. In particular alongside 0.01% MS-222, we test a combination of 0.00025%/ 0.0050% propofol and lidocaine. We choose a concentration of propofol with lidocaine identified as the most successful combination in anaesthetising and recovering adult fish [122, 123]. We chose 2-phenoxyethanol due to its rapid action and fast, uneventful recovery in other fish [124].

Next we determined a protocol for repeated anaesthesia during the juvenile stage. We used the four most promising protocols from experiment B in a repeated anaesthetic trial and monitored the fish for any developmental limitations or reduced survival. We showed that of these four protocols, a combination of 0.085% 2-phenoxyethanol with maintenance by cotton pad was the most suitable protocol for imaging due to low breath rate, although a concentration of 0.01% MS-222 with maintenance by immersion was also suitable. Finally, we repeat the experiment using the most successful protocol increasing the frequency of the anaesthesia to every 2 days instead of every 4 days for 10 fish. This experiment too meets our criteria for success, therefore we propose that this protocol (2-phenoxyethanol at a concentration of 0.085% maintained using a cotton pad doused in anaesthetic) provides the best currently available protocol for repeated anaesthesia with recovery during the 8-16 mm Standard Length (SL) sensitive stages.

7.5. Materials and Methods

7.5.1 Ethics Statement

This study was performed with the approval of the University of Bath ethics committee and in full accordance with the Animals (Scientific Procedures) Act 1986, under Home Office Project Licenses 30/2937 and P87C67227.

7.5.2 Facility conditions

Fish housing

All fish were wild-type AB strain, bred in-house for >20 years. The fish were housed according to FELASA recommendations [125], in tanks filled with circulating system water at 28±0.2oC. System water is made up from Reverse Osmosis water, with synthetic sea salt added at an amount such that the conductivity of the water is kept at approximately 800uS. Average water quality data are as follows: pH: 7.30, general Hardness: 160mg/l CaCO3, Ammonia: 0 mg/l NH3, Nitrite: 0 mg/l NO2, Nitrate: <10 mg/l NO3- and Conductivity: around 800μS/cm. Light cycle is 14 hours light/10 hours dark (Lights on at 08:00, off at 22:00 daily). All fish were in general good health and no specific diseases were observed throughout the colony. Prior to the experiments and in experiment A, fish were housed in groups of no more than 30 fish. From the start of experiments B and C, fish were housed in groups of 10. From the start of experiment D, fish were housed individually.

Feeding

The fish were fed Paramecium from 5-15dpf, followed by Ziegler Larval AP100 powder from day 16 - 22dpf. From 23dpf onwards they are fed Sparos Zebrafeed and Brine Shrimp.

7.5.3 Schedule 1 killing

Fish euthanized by schedule 1 killing were given an overdose of an anaesthetic (0.2% MS-222) followed by crushing of the head to ensure death.

7.5.4 Anaesthetic preparation

MS-222 preparation

A 0.4% buffered MS-222 stock solution was prepared by dissolving ethyl-2-amino-benzoate methanesulfonate powder (Sigma-Aldrich) in system water from the Bath Zebrafish Facility, and the pH adjusted to 7.0 using sodium bicarbonate. Working dilutions were prepared by diluting this stock solution with sterile water.

2-phenoxyethanol preparation

A working solution of 0.085% 2-phenoxyethanol was prepared by diluting 2-phenoxyethanol (Sigma-Aldrich) with system water.

Propofol/lidocaine preparation

A working solution of 0.00025%/0.0050% Propofol/lidocaine mix was prepared by dilution of 1% propofol (Lipuro 1%, B. Braun Melsungen AG, Germany) and 1% lidocaine hydroxide (1%, Braun, Queluz de Baixo, Barcarena, Portugal) with system water, as described in [123]. We note that propofol/lidocaine mixture is a lipid containing emulsion which is not freely soluble in water. Whilst we did not observe visible precipitation when preparing the diluted anaesthetics, we cannot rule out some reduction of the drug concentration in this way.

7.5.5 Experiment A

This experiment aimed to determine the concentration of MS-222 required to successfully anaesthetise and recover juvenile fish, monitoring their subsequent recovery. Here, we anaesthetised batches of 10 to 30 juvenile fish (sized between 7-20mm SL), according to the dish capacity, in a solution of MS-222 at several different concentrations (between 0.008% - 0.02%) for 30 minutes and then recovered them in system water. We chose 30 minutes to reflect the time required for precise orientation and mounting, plus image acquisition when using an epifluorescent or confocal microscope. Animals were randomly allocated from the offspring of more than 10 sets of zebrafish parents.

SL is the distance from the snout to the caudal peduncle. In pre-flexion larvae that do not have a caudal penduncle, SL is the distance from the snout to the posterior tip of the notochord.MS-222 was prepared as described in Section 2.2.1 and placed in a 90 mm petri dish for use. Multiple batches of fish were used to test each concentration. For each, between 10 and 30 zebrafish were collected using a plastic tea-strainer, excess water removed by briefly blotting on paper towel, and then fish transferred to the petri dish with the anaesthetic. After all the fish were transferred, they were monitored over

Figure 7-1: ***Standard length (SL) as described by Parichy et al [29].*** *SL is the distance from the snout to the caudal peduncle. In pre-flexion larvae that do not have a caudal penduncle, SL is the distance from the snout to the posterior tip of the notochord.*

the next 30 minutes. Fish were considered unconscious once they no longer moved freely and were not responsive to being touched gently with a mounted needle. Movement under anaesthetic (MUA) was evaluated three times; at 10, 20 and 30 minutes post immersion. Any fish was counted as MUA if a) the fish did not become anaesthetised within the first 20 minutes of being immersed in MS-222; b) the fish responded to soft touch after 10 minutes of being deemed unconscious (tested by gently touching the lateral side of the fish with forceps); or c) the fish was observed spontaneously moving after being deemed unconscious. Where fish became unconscious, they were moved by pipette to a small volume of anaesthetic (of the same dose) where the standard length (shown in Fig 7-1) was measured using a millimetre ruler, before returning to the petri dish for the rest of the treatment. At the end of the 30 minutes fish were transferred by plastic pipette for recovery in system water. Individuals that were slow to recover spontaneous movement were irrigated with water to aid recovery. Any fish that failed to recover within 30 minutes were noted, their length measured and then euthanized by Schedule 1 killing. Results of fish recovery at certain concentrations informed subsequent testing. For example, if the recovery rate at a particular concentration was too low for fish of a certain length, a higher concentration was never tested on another batch of similar sized fish. Similarly if the MUA was too high at a particular concentration for fish of a certain length, a lower concentration was never tested on another batch of similar sized fish. All fish were anaesthetised and recovered in water at similar temperatures (around 28 degrees Celsius).

7.5.6 Experiment B

This experiment aimed to obtain a workable period of anaesthesia with successful recovery for fish in the sensitive juvenile stages, testing a diverse set of anaesthetics derived from the literature. We distinguished induction - the method used to initially anaesthetise the fish - from maintenance - the method used to maintain the fish under anaesthesia. We anaesthetised 60 fish individually according to one of six protocols (10 fish each) outlined in Table 7.1. We chose the sample size of N=10 in accordance with other similar studies [123, 122] and as a minimal but effective choice. Since we were aiming for 90-100% recovery rate, by using N=10 fish, if the population is 90% successful we can be 95% certain that the recovery rate lies between 70-100%. If the population was 99% successful we could be 95% certain that the recovery rate lies between 94-100%. Informed by the results of experiment A, we reduced the time under anaesthetic to 10 minutes and diversify our anaesthetic type and method. We chose 10

Protocol	Anaesthetic	Dose	Induction	Maintenance	Time maintained (minutes)
1	MS-222	0.01%	Immersion	Immersion	10
2	MS-222	0.01%	Immersion	Cotton pad	10
3	2-phenoxyethanol	0.085%	Immersion	Immersion	10
4	2-phenoxyethanol	0.085%	Immersion	Cotton pad	10
5	Propofol/Lidocaine	0.00025%/ 0.0050%	Immersion	Immersion	10
6	Propofol/Lidocaine	0.00025%/ 0.0050%	Immersion	Cotton pad	10

Table 7.1: *Protocol numbers and the corresponding procedures.*

minutes to reflect the time that might sometimes be required for precise orientation and mounting, plus image acquisition, e.g. using an epifluorescent microscope. Induction in all protocols and maintenance in protocols 1, 3 and 5 was the same as in experiment A: the fish were immersed in a water bath containing the anaesthetic. Fish in protocols 2, 4 and 6 were maintained under anaesthetic by being placed on a cotton pad doused in the anaesthetic. The objective of testing different maintenance methods was to see if a particular method could reduce any negative effects of using a particular anaesthetic e.g. long recovery times or spontaneous movement.Animals were randomly allocated to experimental groups from the offspring of 3 sets of adult zebrafish parents.

At the start of each experiment, the anaesthetic of the correct concentration was prepared (as described in Section 2.2.1) and solutions were poured into a petri dish for use. If the fish was to be maintained by cotton pad, a cotton pad doused in the same anaesthetic would be placed on a second petri dish.

For each protocol, anaesthesia was induced by immersing each fish in the appropriate anaesthetic. Fish that were to be maintained on a cotton pad (protocols 2,4,6) were removed from the anaesthetic using a plastic pipette once they had lost touch response, and placed gently onto the cotton pad. At 10 minutes, the fish were removed from the anaesthetic and recovered in a petri dish of water; where fish were slow to recover spontaneous movements, water was flushed over the gills using a plastic pipette to aid recovery. Fish that did not recover within 30 minutes were recorded and their length measured, before being euthanized by Schedule 1 killing. The mean size of fish in this experiment was 9.99 mm with a standard deviation of 1.4 mm.

In contrast to experiment A, only one fish was anaesthetised in a petri dish at any given time. This allowed us to accurately measure the respiratory rate as well as the time taken for induction of anaesthesia, loss of touch response, recovery of movement and recovery of equilibrium for each individual fish.

Each fish was considered to be anaesthetised once spontaneous movements stopped, and to have lost touch response when it no longer responded to being tapped gently on the fins by a pair of blunt forceps. The respiratory rate was taken once the fish had been under anaesthetic for at least 5 minutes. The rate was measured by counting the number of movements of the mouth made in 60 seconds (i.e. breaths per minute (bpm)) under a stereomicroscope. The heart rate was taken subsequently to the respiratory rate and was measured by counting the number of times the heart pumped blood

Protocol	Anaesthetic	Dose	Induction	Maintenance	Time maintained (minutes)
1	MS-222	0.01%	Immersion	Immersion	10
4	2-phenoxyethanol	0.085%	Immersion	Cotton pad	10
5	Propofol/Lidocaine	0.00025% /0.01%	Immersion	Immersion	10
6	Propofol/Lidocaine	0.00025% /0.01%	Immersion	Cotton pad	10

Table 7.2: ***Protocol numbers and the corresponding procedures.***

in 60 seconds under a stereomicroscope. The standard length (shown in Fig 7-1) was measured using a ruler. The time taken to regain movement (equilibrium) was measured as the time between the fish being placed back into fresh water and moving for the first time (regaining equilibrium and swimming freely).

7.5.7 Experiment C

Building on the results from experiment B, we here aim to identify an optimal protocol for repeated anaesthesia with successful recovery, focusing on protocols 1, 4, 5 and 6. Due to the significantly low recovery rate of protocol 3 in experiment B, we removed protocol 3 from experiment C; furthermore, an initial trial using protocol 2 revealed a high number of fish exhibiting MUA, thus protocol 2 was also removed. A preliminary trial also suggested that the dose of propofol/lidocaine used in Experiment B was too low for successful repeated anaesthesia, and thus the lidocaine dosage was doubled, informed by a recent study [123]. Thus the protocols used in Experiment C along with their corresponding dose, induction and maintenance techniques are given in Table 7.2. For each protocol a batch of 10 fish were anaesthetised every 4 days for 20 days for a total of 6 repeats. Again, we chose the sample size of N=10 in accordance with other similar studies [123, 122] and as a minimal but effective choice, as above. Treated fish were monitored for potential accumulated effects of anaesthetic over this period, focusing on both growth as well as changed sensitivity to the anaesthetic. Animals were randomly allocated to experimental groups from the offspring of 3 sets of adult zebrafish parents.

At the start of the first repeat, all fish in all protocols were aged 21 dpf and the average size of the fish was 5.45 mm SL, 5.4 mm SL, 5.75 mm SL and 5.65 mm SL for protocols 1, 4, 5 and 6 respectively. We anaesthetised the fish in each batch as described in Experiment B. During each repeat we take the same measurements as those outlined in Experiment B. The fish was also checked for any obvious malformations or developmental retardation resulting from the repeated anaeasthesia when the breathing rate was measured. In all cases, there was no obvious malformations or developmental retardation nor any histological changes. After recovery for each repeat (except for the last) the fish were returned to the tank with the rest of the fish from the same experiment group. They remained in the fish facility with free-flowing water and feed until the next repeat. We repeated the test every four days until we had completed 6 repeats and the fish were 41 dpf. We note that in all cases, fish appeared

to be behaving normally between repeated exposures i.e. fish did not move especially quickly or slowly immediately after being returned to the tank, nor 2 days later prior to being anaesthetised again. All fish that survived the six repeats were subsequently euthanized by Schedule 1 killing. One concern was whether or not fish may react to the cumulative effect of the dose i.e. the effect of repeated dosing. A confounding variable for determining whether there were any effects caused by repeated dosing was size; size would increase as the repeat number increased due to growth, and, as we had already found in experiment B, had an effect on susceptibility to the anaesthetic under some conditions. Therefore, in order to control for the effects of SL, at each time point we also performed the same experiment on 5 fish of different lengths that had not been anaesthetised before. Therefore in total we had N=30 fish per protocol ranging in size between 8-16mm SL that were anaesthetised once, and therefore could be used to distinguish effects of repeated dose from effects of size, which we do through using multiple regression analysis.

7.5.8 Experiment D

Building on the results from experiments B and C, we test whether protocol 4, the most successful protocol of experiment C, can be used to repeatedly anaesthetise fish at more frequent intervals. In particular, we test whether juvenile (8mm SL - 16 mm SL) fish are affected by repeated anaesthesia by protocol 4 every 2 days starting at 23 dpf and for a period of 18 days (9 repeats). Animals were randomly allocated to experimental groups from the offspring of 3 sets of adult zebrafish parents. We anaesthetise the fish in each batch as described in Experiment B and C. During each repeat we take the same measurements (except for heart rate per minute and breaths per minute) as those outlined in Experiment B and C. After recovery for each repeat (except for the last) the fish were returned to individual tanks. They remained in the fish facility with free-flowing water and feed until the next repeat. We repeated the test every two days until we had completed all 9 repeats and the fish were 39 dpf. All fish that survived the nine repeats were subsequently euthanized by Schedule 1 killing. Unlike in experiment C, we did not provide a control group for the effects of SL. Since we had already determined that there were minimal effects of repeat number and SL under protocol 4, we deemed it unnecessary to control for size again in this study, thus reducing the number of animals required. As in Experiment C, for all cases, fish appeared to be behaving normally between repeated anaesthetic treatments.

7.5.9 Statistics

Experiment A

To determine the 90% confidence interval for Fig 7-2, we used confint.xls by John C. Pezzullo (jcp12345@gmail.com), which calculates the upper and lower bound of the limit using the Clopper and Pearson method [126] before uploading as a csv file for analysis in python.
The raw data and computed confidence intervals, as well as the code for generating the bargraphs are given at: https://github.com/JenniferOwen/Experiment_A_Data_Analysis_Anaesthetic

Experiments B-D

A normality test was performed for the data in experiments B, C and D, revealing that not all datasets were normally distributed. Therefore the significance test used for all bar plots was the Mann-Whitney-U test. A multi-linear regression was determined using the statistics tool in python for experiments C and D. The code and raw data for each experiment is given as follows:
Experiment B: https://github.com/JenniferOwen/Experiment_B_Data_Analysis_Anaesthetic
Experiment C: https://github.com/JenniferOwen/Experiment_C_Data_Analysis_Anaesthetic
Experiment D: https://github.com/JenniferOwen/Experiment_D_Data_Analysis_Anaesthetic

7.6. Results

7.6.1 Experiment A - longer duration anaesthesia with recovery cannot be readily obtained using MS-222 for juveniles within a sensitive growth range

This experiment aimed to determine the concentration of MS-222 required to anaesthetise successfully juvenile fish, whilst ensuring their subsequent recovery. Our secondary aim was to test what impact, if any, the size of the fish (a proxy for development) has on their sensitivity to MS-222 during this period. We anaesthetised batches of 10 to 30 juvenile fish in solutions of MS-222 at several different concentrations (between 0.008% - 0.02%) for 30 minutes and then recovered them in system water. To assess possible incomplete anaesthesia, we defined Movement Under Anaesthesia (MUA) as any voluntary movement of fish made after 5 minutes of being treated with anaesthetic. If any fish exhibited MUA whilst anaesthetised under any given concentration, then this concentration was deemed too low. Similarly a concentration was considered too high if the survival rate for fish anaesthetised using that concentration was too low (<90%). Fish survived a concentration of MS-222 if after being removed from the petri dish and recovered into fish water, they regained full movement within one hour. Since fish grow significantly bigger during this period (from 8mm-16mm SL), we also measured the standard length ([31]; Fig 7-1) of all fish under anaesthesia and specifically recorded the length of any fish that did not recover.

7.6.2 Length, rather than age, is a better predictor of successful anaesthesia by MS-222

As has been seen previously [31], the length of larvae in each batch showed significant variation, despite all fish being of the same age and raised in the same tank. Moreover, we consistently observed that the effectiveness (ability to successfully anaesthetise and recover) of the MS-222 as an anaesthetic was dependent upon the length of the fish and not necessarily the age (despite these variables being correlated). For example, we would often observe that all fish of a certain length would not survive despite all fish being the same age. For this reason, we focused on testing a wide range of different sized fish, instead of a wide range of ages. Recovery percentage against age is shown in S1 Fig.

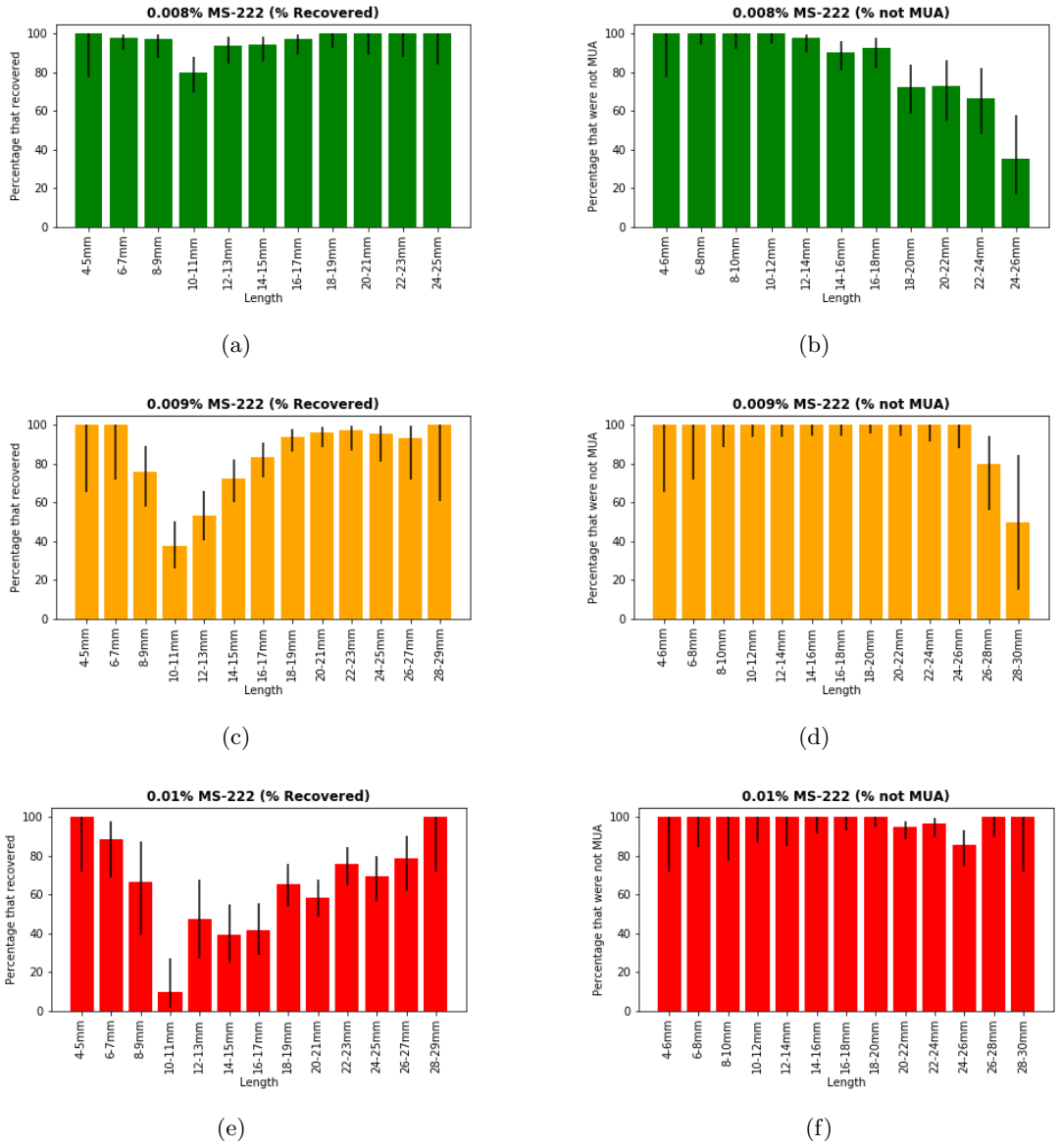


Figure 7-2: ***Metamorphic fish become sensitive to MS-222 around 10mm SL and later become de-sensitised around 24mm SL.*** *(A), (C), (E) Percentage of fish that recovered by length for concentrations of 0.008%, 0.009% and 0.01% MS-222 respectively. (B), (D), (F) Percentage of fish that were not MUA for concentrations of 0.008%, 0.009% and 0.01% MS-222 respectively. Black lines represent the 90% confidence interval for each value, calculated using the sample size (see materials and methods).*

7.6.3 Pre-metamorphic (<6mm SL) and J+ (>16mm SL) fish can be anaesthetised and recovered successfully with MS-222

We found that for fish sized between 4 mm SL and 6 mm SL, MS-222 treatment was very effective for a wide range of concentrations. For example, all fish could be anaesthetised effectively (no MUA) with any concentration from 0.008%-0.01% and >90% recovered successfully (shown in in Fig 7-2A-2E). In fact, pilot studies indicated that fish sized between 4mm - 6mm SL could be anaesthetised and recovered successfully with any concentration up to 0.016%. Anaesthesia was also successfully induced in fish longer than 16 mm SL. For fish sized between 16 mm SL and 26 mm SL, 0.009% MS-222 became the most suitable for effective anaesthesia and good recovery (shown in in Fig 7-2C-2D).). From 26mm SL, a slightly higher dose (0.01% MS-222) became the most suitable for effective anaesthesia and simultaneous good recovery (shown in in Fig 7-2E-2F).

7.6.4 Metamorphic (8 mm SL - 16 mm SL) fish react unpredictably to MS-222

In contrast, fish of intermediate lengths, between 8 - 16 mm SL, were less predictable in their response to MS-222 treatment. For example, when using MS-222 with a concentration of 0.008%, more than 10% of fish did not recover, suggesting the dose was too high (see Fig 7-2A). However, when the dose was lowered to 0.007% MS-222, all fish would remain mobile within the anaesthetic, suggesting the dose was too low. Intriguingly, a small proportion of fish sized between 8 mm SL and 10 mm SL anaesthetised using 0.007% MS-222 would remain mobile throughout the procedure for 20 minutes and then subsequently die in the final 10 minutes, rendering them both MUA and non-recovery. This unpredictability continued to later stages. For example, of fish anaesthetised with 0.008% MS-222 sized between 14 mm SL - 16 mm SL, more than 10% were MUA and more than 10% were non-recovery (see Fig 7-2A and Fig 7-2B), though in this case, these were not the same fish. In general, we were unable to find a concentration of MS-222 that permitted an acceptable frequency of both successful anaesthesia (MUA) and recovery after a 30 min exposure to the anaesthetic.

7.7. Experiment B - Alternative anaesthetics and maintenance methods have potential for repeated anaesthesia for juvenile fish

Consequently, we performed a pilot study of alternative anaesthetics and maintenance techniques, using a diverse set of protocols derived from the literature, with the aim to identify a potential candidate protocol for use in repeated anaesthesia of juvenile fish. In particular, we distinguished induction - the method used to initially anaesthetise the fish - from maintenance - the method used to maintain the fish under anaesthesia. A full set of protocols are described in Table 7.1, but consist of three different anaesthetics, with two different maintenance methods. Fish were anaesthetised for a reduced time of 10 minutes (consistent with a minimal time to image an individual fish) and then recovered in system water. To allow detailed comparison of the protocols, fish were anaesthetised individually, allowing us to monitor a series of key parameters: time taken to induce anaesthesia (i.e. stop movement), time taken to lose touch responsivity, breathing rate, time taken to regain movement post anaesthetic and time taken to fully recover (Fig 7-3). In all cases, there was no MUA in this experiment.

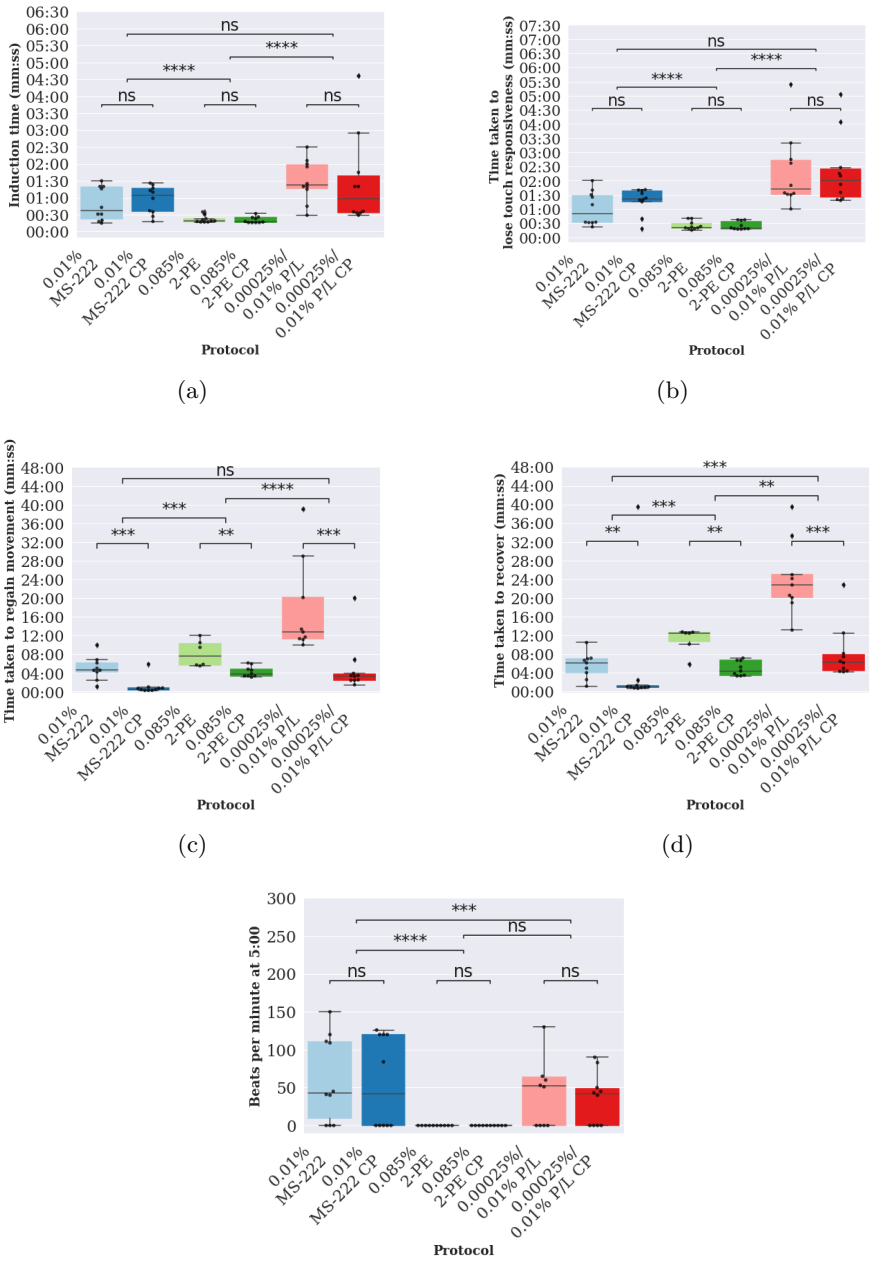


Figure 7-3: ***Protocols 1, 4, 5 and 6 vary in the time taken to induce, lose touch responsiveness, regain movement, regain equilibrium as well as respiratory rate.*** *Boxplots of the time taken to (A) induce anaesthesia, (B) lose touch responsiveness, (C) regain movement post anaesthesia and (D) regain equalibrium post anaesthesia for each protocol 1-6 as defined in Table 7.1. (E) Boxplot of the respiratory rate (breaths per minute - bpm) at 5 minutes for each protocol. Stars indicate significant difference of subsequent repeats as determined using the Mann-Whitney U-test. ns = not significant, '*' indicates p¡0.05, '**' indicates p¡0.01, '***' indicates p ¡ 0.001 and '****' indicates p¡0.0001.*

Protocol	Anaesthetic	Dose	Induction	Maintenance	Survival (%)
1	MS-222	0.01%	Immersion	Immersion	90
2	MS-222	0.01%	Immersion	Cotton pad	100
3	2-phenoxyethanol	0.085%	Immersion	Immersion	50
4	2-phenoxyethanol	0.085%	Immersion	Cotton pad	90
5	Propofol/Lidocaine	0.00025%/ 0.0050%	Immersion	Immersion	90
6	Propofol/Lidocaine	0.00025%/ 0.0050%	Immersion	Cotton pad	100

Table 7.3: ***Protocol numbers and the corresponding procedures with survival percentage***

7.7.1 Anaesthetic choice affects times taken to induce, lose touch response and recover post anaesthetic

Measurements of time taken for induction of anaesthesia and for loss of touch responsiveness are taken prior to the maintenance stage and are thus identical for protocols 1 and 2, 3 and 4 and 5 and 6 respectively. As expected, therefore, there is no statistically significant difference between these protocols as expected (Fig 7-3A and Fig 7-3B). In contrast, the anaesthetic used makes a significant difference. The time taken for loss of touch response was significantly shorter ($p<0.05$) in the case of 2-phenoxyethanol (mean ± sd; immersion: 24 s ± 9 s), than in the case of MS-222 (1m 08 s ± 31 s), and propofol/lidocaine combination (2 m 18 s ± 1 m 17 s) (Fig 7-3A and Fig 7-3B). Furthermore, regardless of the maintenance method, fish anaesthetised using MS-222 regained movement the quickest (immersion: 4m 58 s ±3 m 33, cottonpad: 1m 04 s ± 1 m 42), followed by 2-phenoxyethanol (immersion: 8m 09 s ± 2 m 50, cottonpad: 4m 16 s ± 1 m 09) and then by the propofol/lidocaine combination (immersion: 17 m 35 s ± 10 m 03 s, cottonpad: 4 m 58 s ± 5 m 28; Fig 7-3C and Fig 7-3D).

7.7.2 Maintenance of anaesthesia via cotton pad, as opposed to immersion, improves recovery from anaesthesia

The time taken for recovery from anaesthesia (i.e. to regain movement post anaesthesia and to regain equilibrium) were strongly affected by the maintenance method. Thus, times taken were significantly shorter ($p<0.01$) when the fish was maintained on a cotton pad doused in anaesthetic versus that when it remained immersed in the anaesthetic for the full 10 minutes (Fig 7-3C and Fig 7-3D). We found no significant difference between the respiratory rates of fish maintained on a cotton pad doused in anaesthetic versus that when it remained immersed in the anaesthetic (Fig 7-3E). The mean respiratory rates for each protocol were similar, except that breathing ceased when anaesthetised with protocols 3 or 4. The number of fish that recovered in this pilot study was similar for the two maintenance methods with respect to MS-222 (Table 7.3: 9/10 and 10/10 recovery) and propofol/lidocaine (Table 7.3: 9/10 and 10/10). However, for 2-phenyoxyethanol the number of fish that recovered almost doubled when using a cotton pad for maintenance (Table 7.3: 9/10) over immersion (Table 7.3: 5/10).

Based on the recovery rates, protocols 1,2,4,5 and 6, with 90+% recovery rates, are all suitable protocols for testing repeated anaesthesia. However, for the next stages

Protocol	Anaesthetic	Dose	Induction	Maintenance	Survival (%)	Death occurred on repeat number:
1	MS-222	0.01%	Immersion	Immersion	100	N/A
4	2-phenoxy ethanol	0.085%	Immersion	Cotton pad	100	N/A
5	Propofol/ Lidocaine	0.00025%/ 0.01%	Immersion	Immersion	60	2 (1 fish) 3 (2 fish) 5 (1 fish)
6	Propofol/ Lidocaine	0.00025%/ 0.01%	Immersion	Cotton pad	80	2 (2 fish)

Table 7.4: ***Protocol numbers and the corresponding procedures.***

of testing, we eliminate protocol 2 for repeated anaesthesia for juvenile stages. This is based on observations that cotton-pad maintenance appears to reduce anaesthetic strength seen in experiment B, combined with observations from experiment A that fish become less sensitive to MS-222 during later stages. We also increase the concentrations of lidocaine in protocols 5 and 6 for our repeated anaesthesia trials in order to reduce the time taken to induce anaesthesia.

7.8. Experiment C - Towards a reliable methodology for repeated anaesthesia for juvenile fish

In this experiment we took the subset of protocols determined successful from experiment B, and tested the reliability and impacts of using these protocols for repeated anaesthesia. The full set of protocols are described in Table 7.2, but consist of the same protocols 1 and 4 from experiment B as well as protocols 5 and 6 with a higher concentration of lidocaine. Our aim here was to identify the best candidate protocol for use in repeated anaesthesia and imaging of juvenile fish. The criteria for this was that the protocol had to give consistent low induction and recovery times as well as minimal breathing rate to facilitate imaging. In addition, we assessed whether there were harmful cumulative effects of the anaesthetic trials, such as retarded growth, or inconvenient ones, such as prolonged time taken to induce or recover from, anaesthesia.

Ten fish for each protocol were anaesthetised for 10 minutes and then recovered in system water every 4 days from 21 dpf for a total of 20 days. As in experiment B, fish were anaesthetised individually, allowing us insight into time taken to induce anaesthesia (i.e. stop movement), time taken to lose touch responsivity, heart rate, time taken to regain movement post anaesthetic and time taken to fully recover. We also measured the length of all fish. Since we are interested in imaging the fish, we also measured the breathing rate (visible mouth movements per minute) whilst under anaesthetic, since minimal visible mouth movements would simplify the imaging requirements. We also

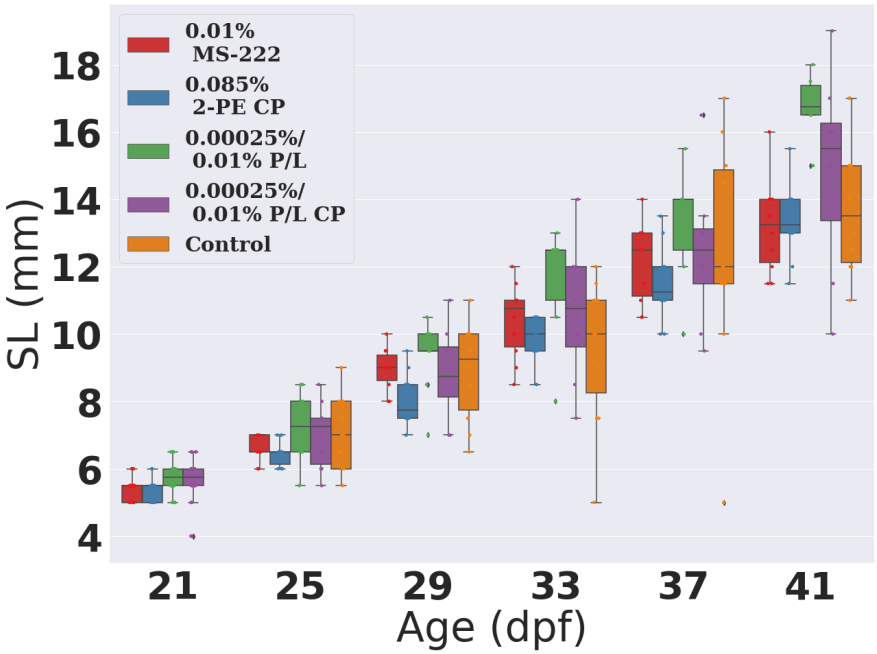


Figure 7-4: ***Growth of the fish is not significantly impacted by repeated dosing using any of the 4 protocols.*** *Barchart with 95% confidence intervals for age vs length (SL (mm)) of fish anaesthetised using protocols 1,4,5 and 6 respectively.*

tested to see if there were any negative effects of repeated treatment such as retarding growth, by comparing the sizes of the fish at each time point with a set of 10 control fish kept in similar conditions. We controlled for whether repeated anaesthesia had effects on individual measurements by simultaneously anaesthetising and measuring a set of fish within the same range of size and age once. Detailed documentation of our full set of results are given in theS2-10 Fig.. An ideal anaesthetic would not be significantly affected by either size or repeat number, but in this section we explore the possibility of cumulative effects of being repeatedly anaesthetised, such as adapting to the dose, or being able to flush it out more readily, or alternatively a cumulative effect leading to overdose, allowing for some effect as long as this was not detrimental. To perform this analysis, for each measurement we generated a multiple linear regression model given in the S1- S4 Tables with variables standard length and repeat number. In Table 7.3 we summarise, for each protocol, the effects of these variables including whether they significantly affect the outcome and if so whether this effect was positive or negative.

7.8.1 Full recovery was observed for protocols 1 and 4 only

The overall recovery rate, that is, the percentage of fish that survived to the end of the full six repeats was 100% for protocols 1 and 4, 80% for protocol 6 and 60% for protocol 5 (see Table 7.4). The fish that did not recover in protocols 5 were all sized within the vulnerable range identified in Experiment A. In contrast, the fish that did not recover following protocol 6 were both very small. They had a mean+s.d. of 6.25±0.75 mm SL.

7.8.2 Repeated anaesthesia did not impact on fish growth

In order to compare the development of the fish over the course of the repeated anaesthesia, we measured the lengths of fish from a control group of 10 fish at each time point. We found that in all protocols for all repeats, the size of the fish in the groups

Protocol 4	Description	Time taken to induce and lose touch responsivity		Visible breaths per minute (bpm)		Time taken to regain movement and recover	
		Repeat number	SL	Repeat number	SL	Repeat number	SL
1	MS-222 0.01%, immersion	++	++	ns	+	-	ns
4	2-phenoxy ethanol 0.085%, cottonpad.	-	-	ns	ns	-	++
5	Propofol/ Lidocaine 0.00025%/ 0.01%, immersion	ns	-	ns	++	ns	-
6	Propofol/ Lidocaine 0.00025%/ 0.01%, cottonpad	ns	-	-	++	ns	-

Table 7.5: ***Susceptibility to anaesthetic treatment under multiple protocols was affected by repeat number and/or standard length.*** *Results based on multiple linear regression models given in the supplementary material. Symbol ++ indicates that the variable causes a significant positive effect and the effect is with a high magnitude i.e. >5 (s or bpm) per repeat number or 5 (s or bpm) per SL. Symbol + indicates a positive effect that is significant, without fulfilling the criteria for ++. ns indicates that the variable is not significant. Symbol - indicates that the variable causes a significant negative effect and the effect is with a high magnitude i.e. >5 (s or bpm) per repeat number or 5 (s or bpm) per SL. Symbol - indicates a negative effect that is significant, without fulfilling the criteria for -.*

were never significantly smaller than the control groups ($p > 0.05$), fish were either statistically similar or were significantly larger than the control, suggesting that growth of the fish is not retarded by the repeated anaesthesia (see Fig 4). Unexpectedly, at the end of the experiment fish from protocol 5 were significantly longer than the controls during later repeat numbers ($p < 0.05$). We attribute this to the effect of attrition of fish from this group for example, from repeat number 4, fish numbers were reduced from 10 to 6. Fish kept in smaller numbers typically consume more and grow more quickly, due to decreased competition for food [127].

7.8.3 Protocols 4 and 6 achieve anaesthesia most rapidly

Since the time taken to become anaesthetised is included within the full anaesthesia time, a good anaesthetic should take a short time (<1 min) to induce anaesthesia and lose touch responsivity. Furthermore, it should not be significantly affected by repeat number and SL, or if it is significantly affected, it should be with only a small magnitude. In Fig 7-4A - 5B, we compare the time taken to induce anaesthesia and time taken to lose touch responsivity for all four protocols, averaged over the 6 repeats.

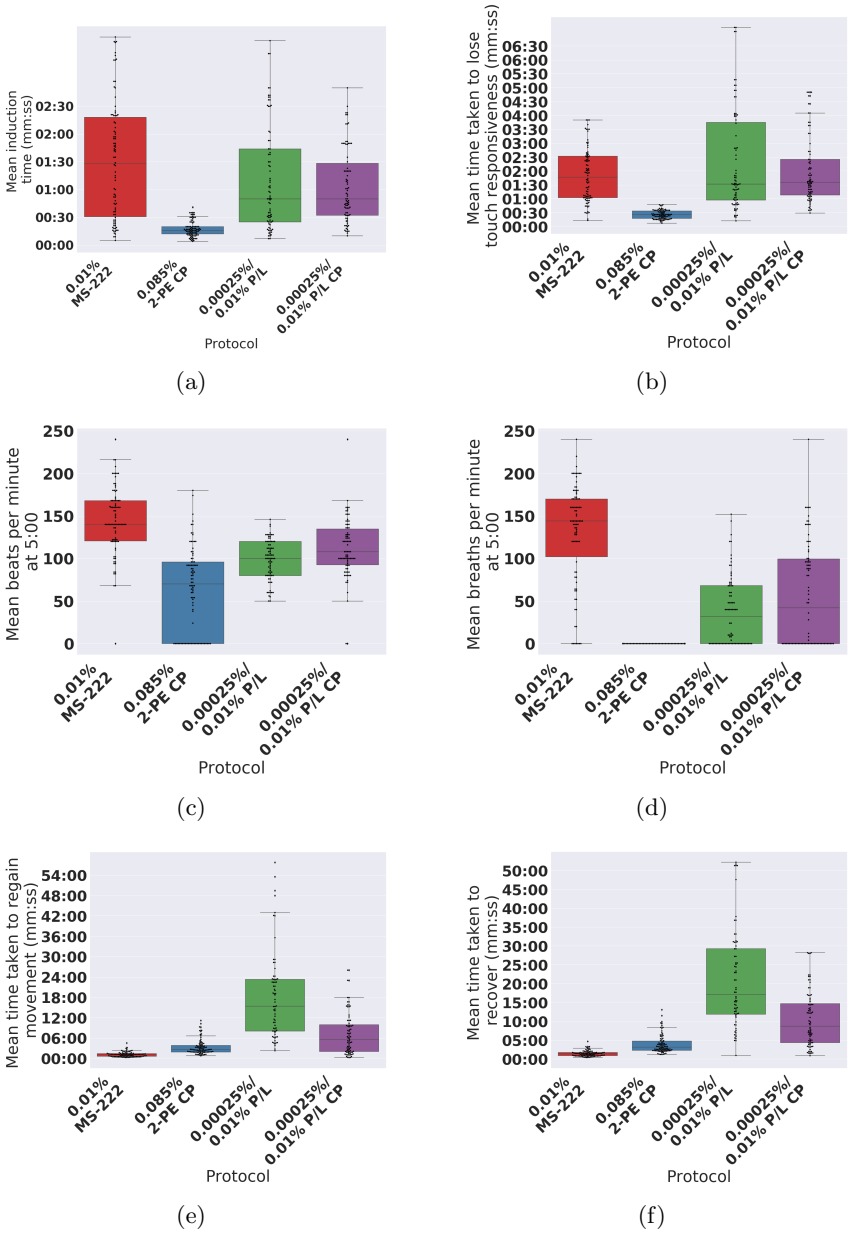


Figure 7-5: ***Protocols 1, 4, 5 and 6 vary in the average time taken to induce, lose touch responsiveness, regain movement, regain equilibrium as well as respiratory rate when repeatedly anaesthetised every 4 days.*** *Barplot of the mean (A) induction time, (B) time taken to lose touch responsivity, (C) heartbeats per minute at 5:00, (D) visible breaths per minute at 5:00, (E) time taken to regain movement and (F) time taken to fully recover against repeat number for each protocol 1,4,5 and 6. Mean taken over all repeats. P/L corresponds to Propofol/Lidocaine, PE corresponds to phenoxyethanol, CP corresponds to Cotton Pad maintenance.*

We found that protocol 4 stands out as being the quickest both to induce anaesthesia and lose touch responsiveness (mean ± sd: 26 s ± 9 s), while all the others are slower for both measures and similar to each other. Protocol 5 is the slowest to lose touch responsiveness (mean ± sd: 1 m 54 s ± 1m 33) and is the least desirable based on time taken to achieve anaesthesia. In Table 7.5 we compare the effects of repeat number and SL on the time taken to induce and lose touch responsivity. We found that fish repeatedly anaesthetised using protocol 1 became quicker to achieve anaesthesia with both repeat number and SL indicating that these factors cause increased susceptibility to the uptake of MS-222. In contrast, fish repeatedly anaesthetised using protocol 4, became slower to achieve anaesthesia with repeat number and SL indicating that these factors cause decreased susceptibility to the uptake of 2-phenoxyethanol, although the magnitude of this effect is not so large. Protocols 5 and 6 were not significantly affected by repeat number, but the time taken to achieve anaesthesia decreased with SL indicating that fish become more susceptible to the propofol and lidocaine combination as the fish develops. Taken together, these data suggest protocol 4 as the most suitable based on the rapidity and reliability of achieving anaesthesia and recovery.

7.8.4 Protocol 4 suppresses breathing rate.

Since we will use our method in order to image fish, a protocol that suppresses breathing rate would be desirable. Furthermore, the extent of this suppression should not be strongly affected by repeat number nor SL. In Fig 7-4D we compare the average breathing rate for different protocols. In contrast to the other three protocols, fish anaesthetised using protocol 4 did not visibly move their mouth during any observations. Fish anaesthetised using protocol 1 displayed the most rapid breathing (mean ± sd: 130 ± 59 breaths per minute). Curiously, fish anaesthetised using protocols 5 and 6 would breath very quickly for short periods of time (≈ 10s) and then would stop for a long time. In Table 7.5 we compare the effects of repeat number and SL on breathing rate. Interestingly, breathing rate in fish anaesthetised using protocol 4 was not significantly affected by SL nor repeat number; indeed breathing rate was consistently 0 over the entire experimental time course. Meanwhile protocols 5 and 6 were significantly and strongly positively affected by SL, and protocol 5 was also significantly and strongly negatively affected by repeat number. Thus, protocol 4 stood out as being best-suited to our aims.

7.8.5 Protocol 4 is most suitable method based on time taken to recover

A good anaesthetic should take a short time (<5 mins) to regain movement and fully recover. Furthermore, it should not be strongly affected by repeat number nor SL. In Fig 7-4E-5F we compare the time taken to regain movement and fully recover for each of the different protocols. Fish anaesthetised using protocol 1 recovered the quickest (mean ± sd: 1 m 24 s ± 46s), followed by protocol 4 (mean ± sd: 4 m 04 s ± 2 m 31 s) and thus are the most suitable by this measure. In contrast fish anaesthetised under protocols 5 and 6 took a very long time to recover (between 15 m - 30 m). In Table 7.3 we compare the effects of repeat number and SL on the time taken to regain movement and recover. Interestingly all protocols were significantly affected by one or both of SL and or repeat number with a high magnitude. Fish anaesthetised using protocol 1 were negatively affected by repeat number but not SL, suggesting that repeated dosing may help the fish to flush MS-222 from their system. In contrast, fish anaesthetised using

Protocol	1	4	5	6
Average time taken to induce and lose touch responsivity quick (<1m)	X	✓	X	X
Time taken to induce touch responsivity not affected by SL and repeat number with a high magnitude	X	✓	✓	X
Breaths per minute <10	X	✓	X	X
Breaths per minute not affected by SL and repeat number with a high magnitude	✓	✓	X	X
Time taken to recover and regain movement and recover (<5m)	✓	✓	X	X
Time taken to recover and regain movement not effected by SL and repeat number with a high magnitude	X	X	X	X
Growth not retarded by repeated dose	✓	✓	✓	✓
100% recovery	✓	✓	X	X
Total score	4/8	7/8	2/8	1/8

Table 7.6: ***Summary table of protocols and successes.***

protocols 5 and 6 were negatively affected by repeat number but not SL, indicating that as fish develop they may become more efficient at removing the propofol/lidocaine mix from their system. Interestingly the time taken for fish anaesthetised by protocol 4 to recover was negatively affected by repeat number and positively affected by SL at a similar magnitude. Since between each repeat fish grew on average 1mm, these effects essentially cancel out. Therefore, protocol 4 is the most suitable protocol for repeated anaesthesia, based on the time taken to recover.

7.8.6 Maintenance of anaesthesia via cotton pad, as opposed to immersion, improves recovery and time taken to recover from anaesthesia for propofol/lidocaine mix

Protocols 5 and 6 both use a propofol/lidocaine mix described in Table 7.2, except in protocol 5 we maintain fish in anaesthetic solution, whereas for protocol 6 we maintain fish on a cotton pad soaked in anaesthetic. We found that the time taken to recover is approximately halved using the cotton pad versus maintenance in anaesthetic (mean ± sd: 9 m 07 s ± 6 m 03 versus 21 m 19 s ± 15 m 16 s) suggesting that the maintenance method of a cotton pad helps to reduce the effects of the anaesthetic during the recovery period. Consistent with this, we also found that the recovery rate was higher for protocol 6 over protocol 5.

7.8.7 Protocol 4 is the most suitable protocol for repeated anaesthesia.

In Table 7.6 we compare the different protocols by the success measures given in this section. Meeting 7/8 of the criteria, we deem that protocol 4 is substantially the most suitable for repeated anaesthesia under the conditions tested. Protocol 1 is second, meeting 4/8 criteria, and would be suitable for repeated anaesthesia that did not require imaging. In contrast, we deem that protocols 5 and 6 are not suitable for repeated anaesthesia during the juvenile period tested. Fish anaesthetised using these protocols took a long time to be anaesthetised and recover, visibly breathed frequently and there was not 100% recovery.

7.9. Experiment D - Protocol 4 is effective for frequent, repeated anaesthesia with recovery throughout the susceptible juvenile period

Finally, we performed a focussed study to assess the impact of frequent, repeated anaesthesia with recovery using protocol 4. A batch of 10 fish were repeatedly anaesthetised every 2 days, starting from 23 dpf for a total of 9 repeats (18 days). Unlike in experiment C wherein all 10 fish were housed in one tank throughout the experiment due to space constraints, we housed the fish individually, allowing us to track individual development and anaesthesia effect on fish individually. We note that there was no significant difference between the measured effects of individual fish. As in experiment C, we recorded the time taken to induce anaesthesia (i.e. stop movement), time taken to lose touch responsivity, time taken to regain movement post anaesthetic and time taken to fully recover, as well as the length of the fish. We did not measure heart or breathing rate. We tested whether fish experience retarded growth over the repeats by comparing the sizes of the fish at the start and end point with a set of 10 control fish also housed individually. As we had already established that there were no adverse effects of repeated anaesthetic using protocol 4 in experiment C, in order to minimise the number of fish needing to be involved in this trial we did not control for the effects of repeated versus single anaesthesia. The aim of this experiment was to identify whether protocol 4 was suitable for repeated anaesthesia with more frequent dosing, i.e. that the fish could be anaesthetised and recovered quickly, did not breathe frequently, that there was no MUA and all fish would be recovered, whilst also avoiding developmental retardation.

7.9.1 Growth of the fish was not retarded by the repeated anaesthesia at a higher frequency

In order to compare the development of the fish over the course of the repeated anaesthesia, at the start of the experiment we separated 20 fish from an initial batch of 50 fish with the same age and similar size that had been kept in a tank together since 5 dpf. Of these 20 fish, 10 fish were chosen (uniformly at random) to occupy a control group and the rest were allocated to the experimental group. All fish were housed in individual tanks during the 20 day period. Fish in the experimental group were measured during every anaesthetic trial. Fish in the control group were measured once, at the experimental end point, after 18 days. The mean $\pm$ s.d.length of the experimental and control fish at 31 dpf showed no significant difference (experimental, 16.35 $\pm$ 2.11 mm SL; control, 17.45 $\pm$ 1.9 mm SL.

Protocol 4	Description	Time taken to induce and lose touch responsivity		Visible breaths per minute (bpm)		Time taken to regain movement and recover	
		Repeat number	SL	Repeat number	SL	Repeat number	SL
Every 4 days	-	-	ns	ns	++	-	++
Every 2 days	-	+	ns	ns	ns	ns	ns

Table 7.7: ***Comparison of the significance of repeat number and standard length on the time taken to induce and lose touch responsivity, breathing rate and time taken to regain movement and recover for protocol 4 when used every 2 days vs every 4 days.*** *Results based on multiple linear regression models given in the supplementary material. Symbol ++ indicates that the variable causes a significant positive effect and the effect is with a high magnitude i.e. >5 (s or bpm) per repeat number or 5 (s or bpm) per SL. Symbol + indicates a positive effect that is significant, without fulfilling the criteria for ++. ns indicates that the variable is not significant. Symbol - indicates that the variable causes a significant negative effect and the effect is with a high magnitude i.e. >5 (s or bpm) per repeat number or 5 (s or bpm) per SL. Symbol - indicates a negative effect that is significant, without fulfilling the criteria for -.*

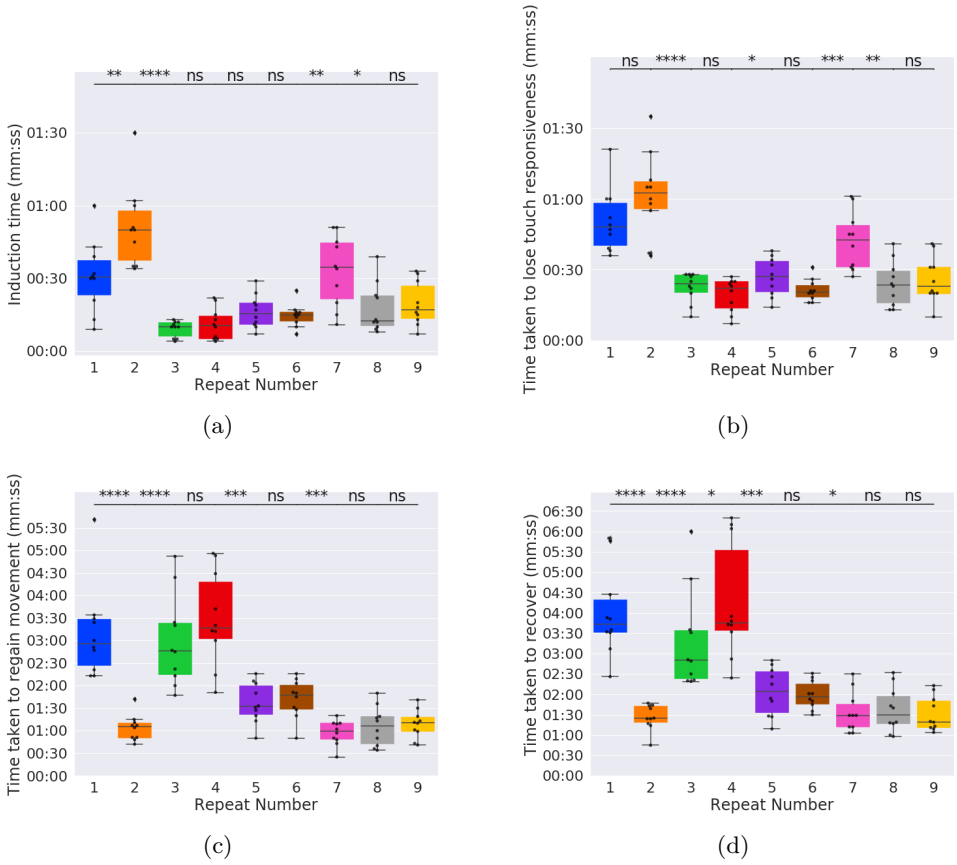


Figure 7-6: ***The time taken to induce, lose touch responsiveness, regain movement, regain equilibrium as well as respiratory rate varies within an acceptable range for fish repeatedly anaesthetised using protocol 4 every 2 days.*** *Boxplot of (A) induction time, (B) time taken to lose touch responsiveness, (C) time taken to regain movement and (D) time taken to recover. Stars indicate significant difference of subsequent repeats as determined using the Mann-Whitney U test. ns = not significant, * indicates p<0.05, ** indicates p<0.01, *** indicates p <0.001 and **** indicates p<0.0001.*

7.9.2 Protocol 4 remained an efficient and effective method for inducing anaesthesia with recovery when used frequently.

We found that increased frequency of inducing anaesthesia showed no significant impact on the suitability of protocol 4 Just as when fish were repeatedly anaesthetised every 4 days, the time taken to induce anaesthesia ($\approx$ 1 min), and to recover from anaesthesia remained low ($\approx$ 5 min see Fig 7-5A-6B). Breathing rate was consistently very low. Finally, we compare the effects of SL and repeat number on the time taken to induce and the time taken to recover for fish repeatedly anaesthetised using protocol 4 (Table 7.7). Interestingly, when the fish were repeatedly anaesthetised every 2 days, SL had a positive effect rather than a negative effect on the time taken to induce anaesthesia. Furthermore, unlike when fish were anaesthetised every 4 days, the time taken to regain movement and recover was not significantly affected by either repeat number of SL.

7.9.3 Protocol 4 is a suitable protocol for repeated, frequent anaesthesia.

Based on these new results, in Table 7.8 we assess protocol 4 using the same criteria as for experiment C. We observe that protocol 4 is successful in fulfilling all criteria when used for repeated anaesthetic every 2 days. Therefore we conclude that protocol 4 is a suitable protocol for frequent repeated anaesthesia during juvenile periods.

7.10. Discussion

Whilst there have been some studies of anaesthesia protocols as applied to adult and larval zebrafish [128, 116, 123, 122], there have been none to our knowledge that have tested anaesthetics during the juvenile period.

Zebrafish larvae and adult fish vary greatly in their response to anaesthetics due to their significantly different physiologies [129, 130]. Adult zebrafish rely on their gills for oxygen absorption. The usual cause of death for adult fish under anaesthesia is by asphyxiation, caused by the blockage of gill ventilation leading to hypoxemia [131, 132]. Deeply anaesthetised adult fish can be kept alive for hours to days, but only if their gills are artificially ventilated [133]. Zebrafish larvae do not have gills and thus rely on cutaneous gas exchange for oxygen absorption. Zebrafish larvae, as a result, are considerably more tolerant to the lethal effects of MS-222 than adult fish.

Anaesthetic treatment for juvenile stages therefore holds uncertainty. During the juvenile period, metamorphosis includes substantial changes in anatomy and physiology, including of the gills. These start to develop between 5 dpf and 33 dpf [129]. Zebrafish switch from absorbing oxygen by cutaneous gas exchange [134], to gill absorption at this stage. Therefore, it is logical to expect a change in tolerance to the lethal effects of anaesthetics during this period. Therefore our aim was to refine the method for anaesthetising juvenile fish to maximise the welfare during this poorly characterised period of development. Initially, we studied the effect of different concentrations of the most commonly used fish anaesthetic: MS-222 on the induction of anaesthesia for 30 minutes exposure and subsequent recovery of zebrafish at juvenile stages. We found that we were unable to determine any dose for zebrafish between the ages of 21 dpf and 60 dpf, such that >90% fish could be both anaesthetised for 30 minutes and also be successfully recovered. We found that survival of the fish for a given concentration of MS-222 varied with the length of the zebrafish, confirming earlier observations that this provided a more accurate measure of developmental stage than age [135]. We

Protocol	**Every 4 days**	**Every 2 days**
Average time taken to induce and lose touch responsivity quick (<1m)	✓	✓
Time taken to induce touch responsivity not affected by SL and repeat number with a high magnitude	✓	✓
Breaths per minute <10	✓	✓
Breaths per minute not affected by SL and repeat number with a high magnitude	✓	✓
Time taken to recover and regain movement and recover (<5m)	✓	✓
Time taken to recover and regain movement not effected by SL and repeat number with a high magnitude	X	✓
Growth not retarded by repeated dose	✓	✓
100% recovery	✓	✓
Total score	7/8	8/8

Table 7.8: ***Summary table of successes for protocol 4 for two different frequencies.*** *In both cases, protocol 4 has a high score and thus is a suitable protocol for repeated anaesthesia.*

identified a critical period related to size, between 8 - 16 mm SL (loosely corresponding to approx. 25 - 45 dpf) when zebrafish experience enhanced sensitivity to MS-222. During this period, fish would remain active in a dose of 0.007% MS-222 and yet were vulnerable to being euthanized by a dose of 0.008% MS-222. In some cases, fish would move spontaneously for up to 20 minutes in the anaesthetic, stop moving abruptly and then die a few minutes later. We have subsequently identified one other report that noted that zebrafish aged between 21-70 dpf are particularly sensitive to MS-222 [111]. We are unsure of the precise cause of death for fish during this period. One possible reason could be linked to gill development during this period. At an approximately equivalent stage (23 - 33 dpf), gill filaments transform from a basic branched shape to a structure close to their final definitive form [129]. During this time, the overall contribution of cutaneous gas exchange to oxygen uptake decreases and the gills become the more dominant site for gas intake. It is possible that the changes which occur during this period leave the fish more vulnerable to the lethal effects of MS-222 by asphyxiation. We next aimed to determine a suitable anaesthesia and recovery protocol for this period. We reduced the duration of anaesthesia required to 10 mins. Furthermore, we broadened our anaesthetic trial to include two other anaesthetics that have received attention recently by the zebrafish community as well as a medium 0.01% dose of MS-222. The combination of propofol and lidocaine was recently tested and recommended for adult zebrafish in a study by Valentim [123]. Propofol is short-acting, rapidly metabolized and less prone to cumulative effects than MS-222 [128]. It is generally considered a very safe drug in other vertebrates and mammals and is commonly used in veterinary medicine, though not often in fish. It has a biological half-life in rainbow trout of 1.1 hours at 17^oC [136] (the estimated plasma half-life of MS-222 is between 1.5 hours and 4.0 hours [137]). Propofol is combined with lidocaine to achieve analgesia. The combination also allows for a decrease in the propofol dose, resulting in a safer analgesia and anaesthesia. We also included a dose of 0.085% 2-phenoxythanol. 2-Phenoxyethanol is commonly used in aquaculture [116] and shows a rapid effect and recovery time. It has a reduced physiological effect when compared with MS-222 and has been recommended in some literature as a suitable alternative to MS-222 for anaesthesia of adult fish [119]. In adult fish, 2-phenoxyethanol is absorbed through the gills and transported via the arterial blood to the CNS. It is rapidly excreted (via the gills) and has a biological half-life in rainbow trout of less than 30 minutes. 2-phenoxyethanol is considered suitable for aquaculture because of its ease of preparation, low cost, rapid action and fast, uneventful recovery in other fish [124]. When left for extensive periods of 96 - 140 hours, 2-phenoxyethanol has been shown to be toxic for zebrafish embryos (<5 dpf, 140h LC50: 461.52-521.55 mg L^{-1}) and fish sized between 30 $\pm$ 5 mm SL (96h LC50: 312.10-349.02 mg L^{-1}) [117]. 2-phenoxyethanol has not been tested for juvenile stages between 8 m SL -16 mm SL and not for the much shorter times considered here. It should be noted that health concerns have been raised about repeated exposure to this compound on the handler [138]; these would suggest that caution should be exercised and specifically that fish at other stages of development should be treated with MS-222 as is current standard practise. For each anaesthetic, we tested two different methods of anaesthesia maintenance. Our aim was to determine the best anaesthetic procedures for juvenile fish, with the plan to take those forward to the next stage of testing consisting of repeated anaesthesia. We found that for all protocols, all fish were successfully anaesthetised and once anaesthetised, remained non-touch responsive (non MUA) throughout the procedure. This suggests that all the anaesthetics trialled in experiment B were appropriate for successful anaesthesia in juvenile zebrafish.

Of the three anaesthetics, 2-phenoxyethanol had the lowest survival rate - across

the two methods, 30% of the fish did not recover post-anaesthesia. The time taken to induce and lose touch responsiveness for these fish was significantly shorter than for fish anaesthetised using MS-222 and the propofol/lidocaine combination. This suggests that 2-phenoxyethanol has a stronger effect on juvenile zebrafish than the other two anaesthetics. Fish anaesthetised using the propofol/lidocaine combination were the slowest to be anaesthetised and the slowest to recover. This suggests that the combination is slow acting yet strongly perjuring for juvenile fish. Our recorded times for this protocol were similar to those observed by Valentim [123] when tested in adult zebrafish. Fish anaesthetised using MS-222 were the quickest to recover. We found the reduction in time spent under anaesthetic - from 30 minutes to 10 minutes - greatly increased the probability of recovery. In experiment A, the proportion of fish sized between 8 mm SL and 16 mm SL that recovered from 30 minutes of 0.01% MS-222 concentration varied from 10% (10 mm SL - 12 mm SL) to 70% (8 mm - 10 mm SL). When the time spent under anaesthesia was reduced to 10 minutes, as in experiment B protocol 1, we observed a high recovery rate of 90% for fish sized between 8 mm SL and 16 mm SL. This suggests that a shorter anaesthetic time is preferential for juvenile fish.

Consistently, we found that fish maintained on a cotton pad were quicker to regain movement and recover equilibrium, compared to fish that were kept in the bath of anaesthesia. This is likely due to a smaller surface area in contact with the anaesthetic. Fish anaesthetised by 2-phenoxyethanol had an increased recovery rate from 50% to 90% when maintained on a cotton pad, suggesting that maintenance by cotton pad can increase non-lethal anaesthesia time for strong anaesthetics. However, another reason could be related to the oxygen availability. One study suggests that the pectoral fins play an important role in mixing up the oxygen boundary layer and hence maintaining high oxygen levels at the skin surface. For example, when early zebrafish larvae aged between 4-20 dpf were anaesthetised to briefly halt fin and gill ventilation, the oxygen concentration in the water near the fins decreased by as much as 50% [139]. Therefore, cotton pad maintenance may play a role in maintaining a high oxygen level at the skin surface at stages when fish are unable achieve efficient gas exchange using their gills. It is not clear whether there are any welfare implications of using one method over another. The cotton pad method could in theory be stressful for the fish as they are outside of their natural environment, however, we found no evidence that would indicate that the animals were more stressed. For example, there was no increase in the breathing or heart rate, known signs of stress in fish [140], using the cotton pad method. We note that if our prediction that the cotton pad method reduces fatality by allowing a better oxygen supply to the fish, an equivalent method would be to partially submerge the fish in the anaesthetic. This method may be more suitable for some imaging purposes, e.g. with transmitted light.

In experiment C, we aimed to determine whether cumulative exposure of remaining methods may produce any negative side effects due to accumulation of the anaesthetic.

Previous investigation of repeated use of MS-222 to anaesthetise fish is limited and does not encompass zebrafish. However, such investigations have suggested that repeated exposure to MS-222 could inhibit development in some fish. For example, when hybrid tilapias were exposed weekly to MS-222, the fish displayed significant reductions in development upon the third exposure and thereafter [141], suggesting that repeated doses of MS-222 may accumulate. Consistent with this, Atlantic salmon repeatedly anaesthetized with MS-222 exhibited upregulation of osmoregulatory genes in the gill and effects on blood parameters [142]. MS-222 anaesthetised rainbow trout have also been shown to ingest 15-20% less food than non-anaesthetised fish for up to 48 hours

post anaesthesia [115]. Food plays an important role in timely development. For maximum weight gain during juvenile development, protein intake is recommended to be 14mg/g average body weight / day [123]. If MS-222 also inhibits feeding in juvenile fish, repeated exposure could greatly stunt food intake and hence arrest development. There is currently no available research into the effects of repeated exposure to propofol/lidocaine combination nor of 2-phenoxyethanol in fish. Despite these concerns, we found that for all protocols the fish grew at a rate comparable to our control group, suggesting that repeated anaesthesia once every 4 days has no impact on development. We found that MS-222 maintained via immersion is suitable for repeated anaesthesia of juvenile fish when used at a concentration of 0.01% for 10 minutes every 4 days. This protocol had a 100% survival rate, did not noticeably impact development and showed a very small effect of de-sensitisation to the anaesthetic over many repeats.

2-Phenoxyethanol maintained on a cotton pad is also suitable for repeated anaesthesia of juvenile fish when used at a concentration of 0.085% for 10 minutes every 4 days. This protocol had a 100% survival rate, did not impact development and showed a very small effect of de-sensitisation to the anaesthetic over many repeats. Repeated exposure to 2-phenoxyethanol has been shown to affect the anaesthetics pharmacokinetics in juvenile Angelfish, causing some tolerance; , it is not clear whether other fish species may also develop tolerance to 2-phenoxyethanol. 2-phenoxyethanol has been shown to cause a stress response and immunodepression in adult seabream [143]. We found no obvious signs that the zebrafish was stressed during the procedure - no rapid breathing for example. An advantage of this method was the very low breath count, making it highly suitable for technique that require fish to stop moving completely, such as repeated imaging. Therefore, out of protocols 1 and 4, protocol 4 was the most suitable for repeated anaesthesia for imaging, due to fish exhibiting a suppression of breathing movements with this method.

In contrast, we conclude that protocols 5 and 6 were not suitable for repeated anaesthesia during this sensitive period. Firstly, under these protocols we observed a less than 100% recovery. Furthermore, fish were vibration responsive throughout the experiments, making them inappropriate for repeated anaesthesia for imaging. Also, unlike fish anesthetised under protocols 1 and 4, which either did not breath during anaesthesia (protocol 4), or did so at a controlled and constant pace (protocol 1), fish anaesthetised under protocols 5 and 6 alternated between not moving their mouth, to opening and closing their mouth rapidly. Further work is required to determine the cause of these differences. One reason could be a lack of oxygen. Opening and closing of the mouth, as well as gill movement is part of the respiratory cycle [144]. Rapid gill movements in zebrafish larvae, is associated with hypoxia [139].

Finally in Experiment D, we showed that protocol 4 remained suitable for repeated anaesthesia with recovery, even when the frequency of anaesthesia was increased from every 4 days to every 2 days. From this experiment we deduce that protocol 4 is suitable for frequent repeated anaesthesia during the juvenile period. Conclusions

In conclusion, we have shown that during the juvenile period (between 6 mm SL and 30 mm SL), the MS-222 dose (administered via water bath) required to successfully anaesthetise zebrafish is dependent on the length of the fish. Moreover, for some sensitive periods (8 mm SL - 16 mm SL) there is no suitable standard dose which can successfully anaesthetise for 30 minutes with both no MUA and full recovery. In contrast, zebrafish sized between 8 mm SL and 16 mm SL can be successfully anaesthetised for 10 minutes in a water bath using MS-222 at a concentration of 0.01% and maintained by immersion (protocol 1). Fish of this size can also be successfully anaesthetised using 2-phenoxyethanol at a concentration of 0.085% when maintained

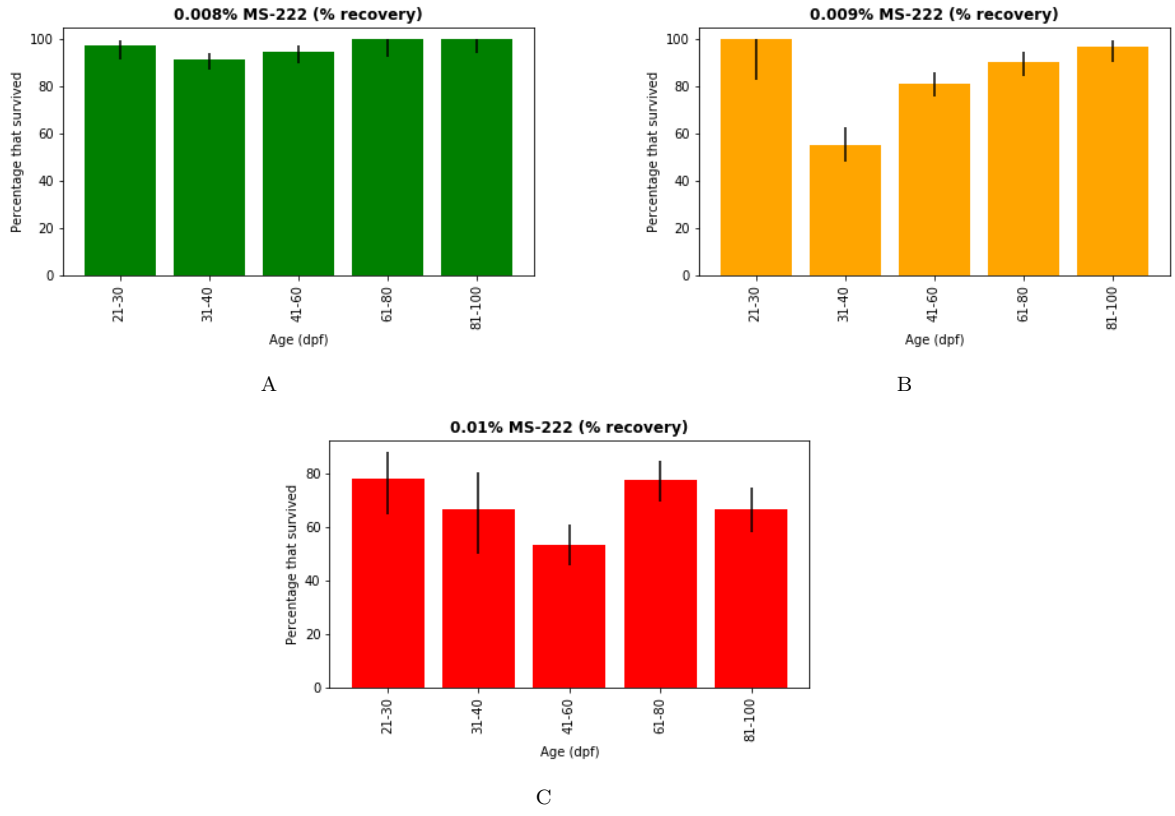
on an anaesthetic doused cotton pad (protocol 4). Protocols 1 and 4 are also successful methods (i.e. 100% recovery, low induction time and recovery times) when used to repeatedly anaesthetise juvenile fish every 4 days (for 6 repeats). Furthermore, repeated anaesthesia of zebrafish over this developmental period does not necessarily effect development. Of the two protocols, protocol 4 is a better protocol for imaging as fish do not visibly move their mouth during anaesthesia. Furthermore, protocol 4 is a suitable method for repeatedly anaesthetising fish when the frequency is increased to every 2 days (for 9 repeats).

Therefore, based on our data presented here, we propose that 2-phenoxyethanol at a concentration of 0.085% maintained using a cotton pad doused in anaesthetic provides the best currently available protocol for repeated anaesthesia with recovery during the 8-16 mm SL sensitive stages, offering a refinement over other adult-focussed protocols.

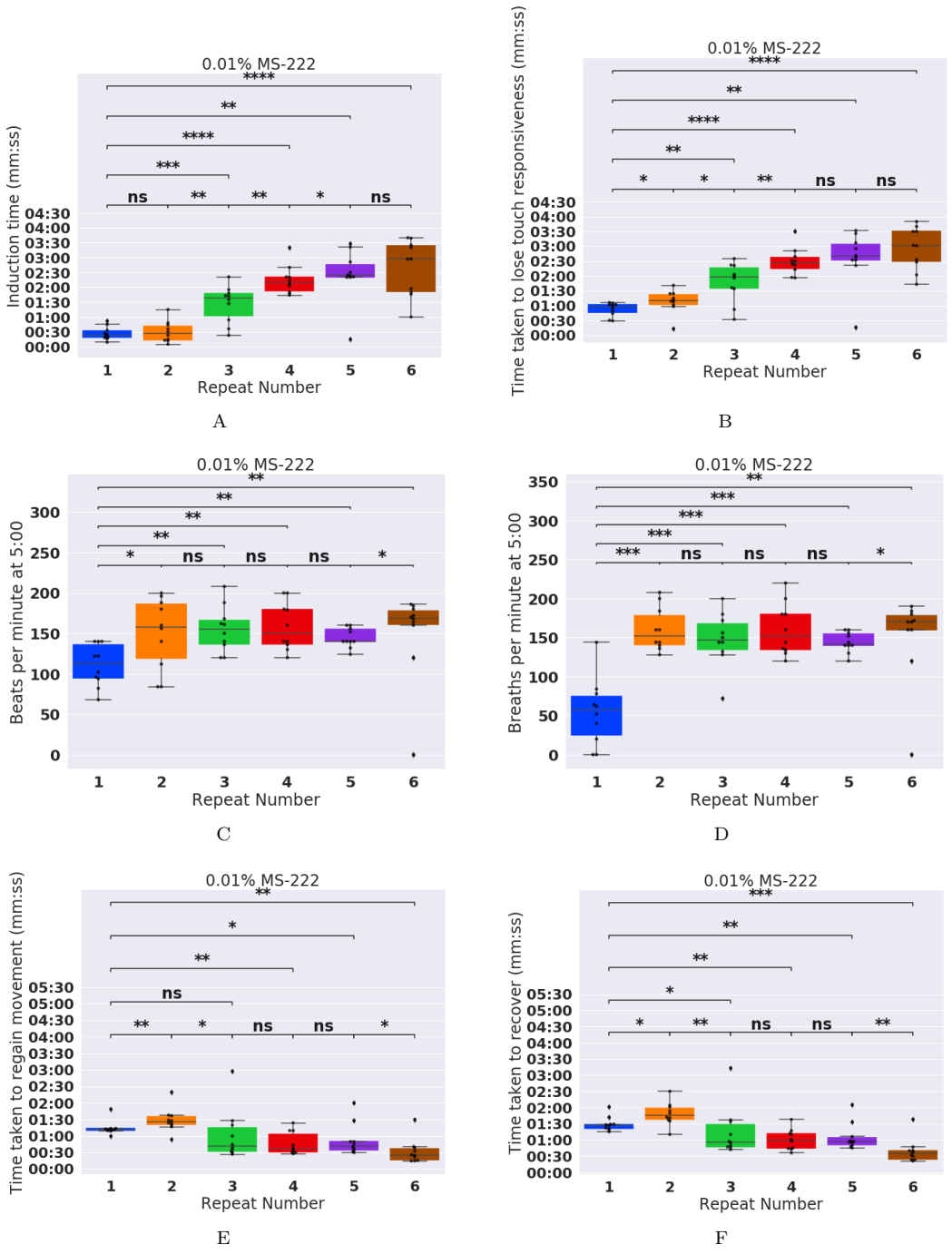
Supplementary Figures to accompany the paper 'A suitable anaesthetic protocol for metamorphic zebrafish'

Jennifer P. Owen, Robert N. Kelsh

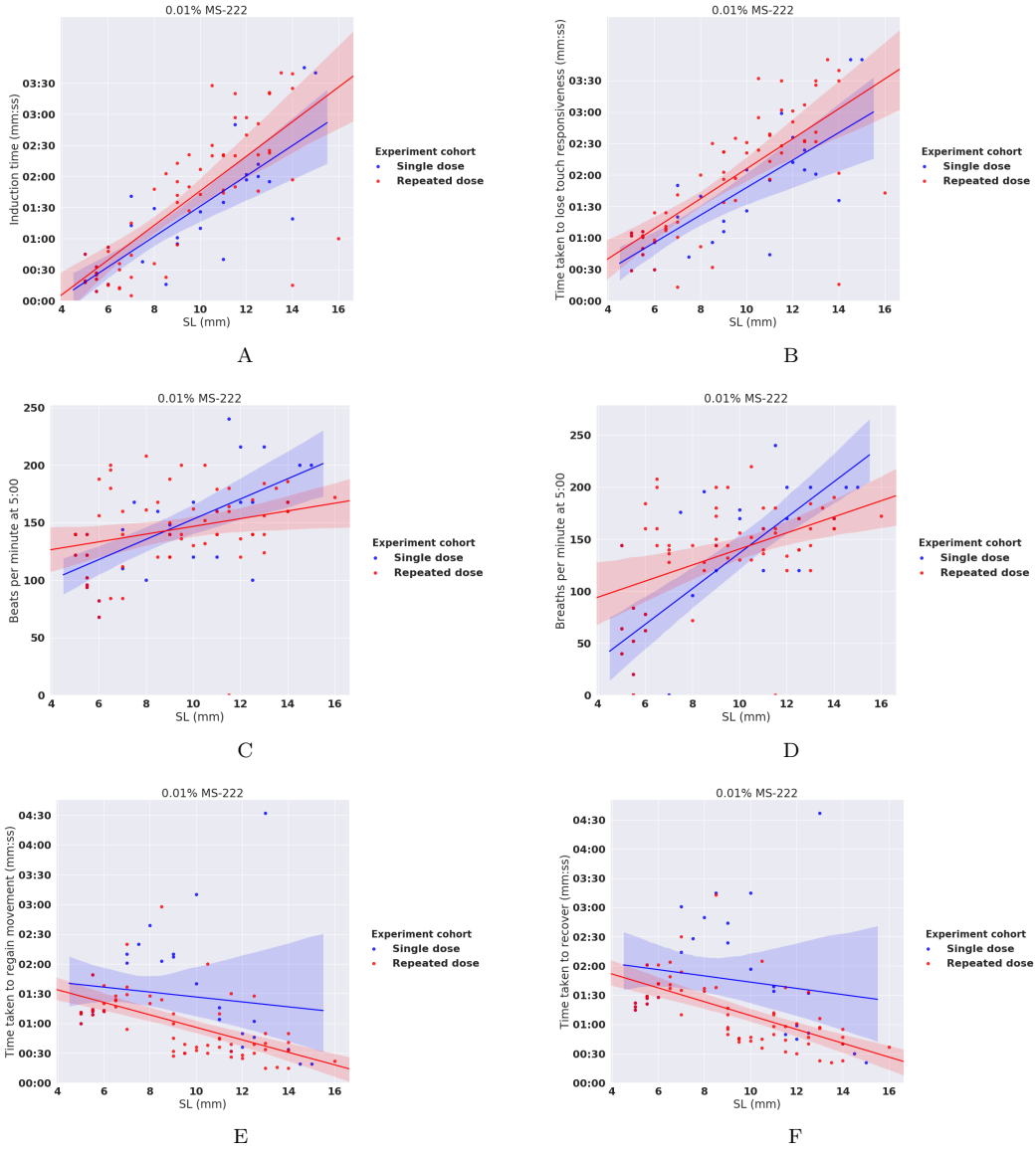
January 3, 2021



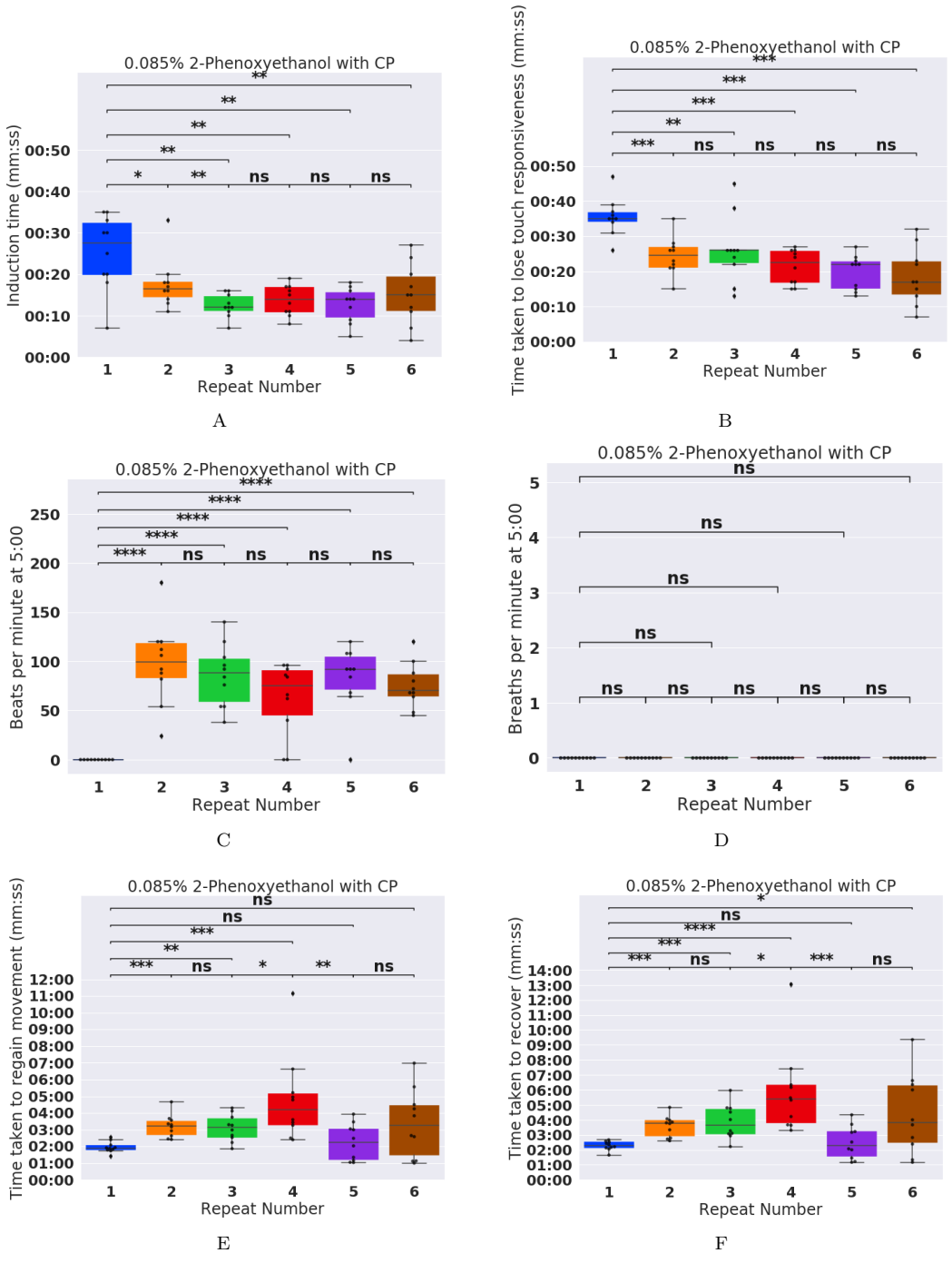
Supplementary Fig. 1: Age is not a good predictor of recovery. (A), (B), (C) Percentage of fish that recovered by age for concentrations of 0.008%, 0.009% and 0.01% MS-222 respectively. Black lines represent the 90% confidence interval for each value, calculated using the sample size (see materials and methods).



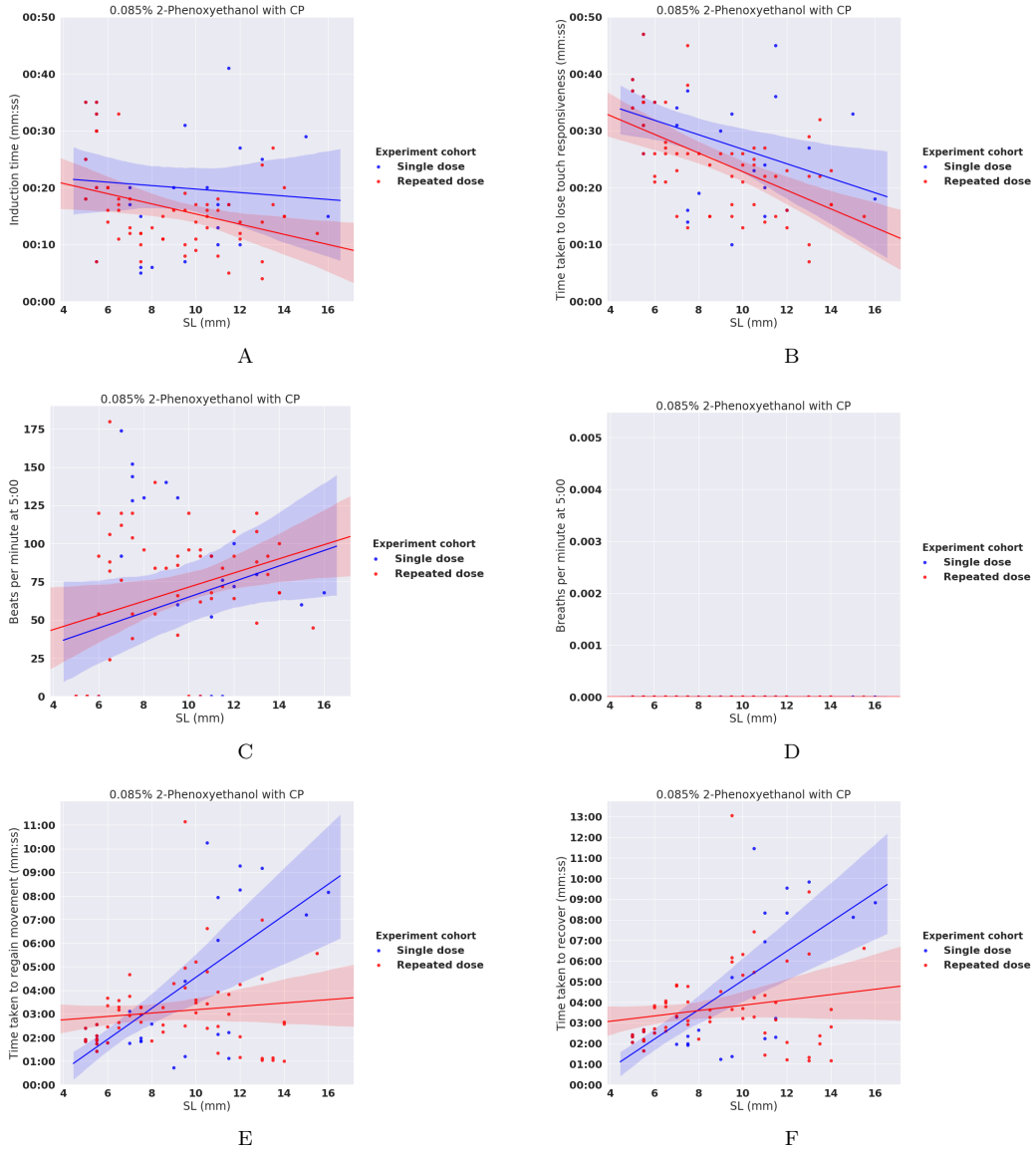
Supplementary Fig. 2: The time taken to induce, lose touch responsiveness, regain movement, recover as well as respiratory rate varies with repeat number when repeatedly anaesthetised using protocol 1 every 4 days. Boxplot of (A) induction time, (B) time taken to lose touch responsivity, (C) heartbeats per minute at 5:00, (D) visible breaths per minute at 5:00, (D) time taken to regain movement and (E) time taken to fully recover against repeat number when using Protocol 1 to repeatedly anaesthetise and recover fish. Stars indicate significant difference as determined using the Mann-Whitney U test. ns = not significant, '*' indicates $p<0.05$, '**' indicates $p<0.01$, '***' indicates $p < 0.001$ and '****' indicates $p<0.0001$.



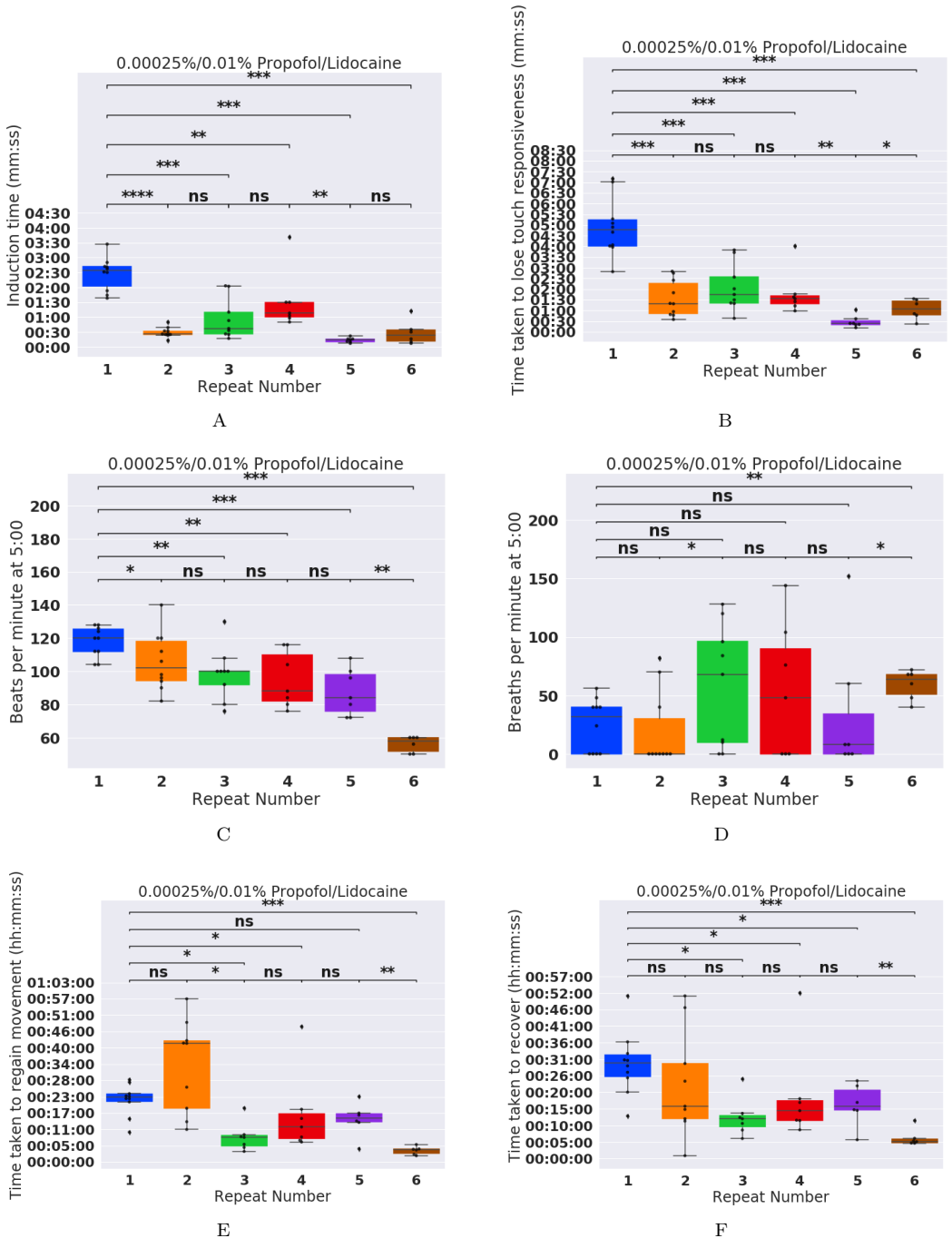
Supplementary Fig. 3: The time taken to induce, lose touch responsiveness, regain movement, recover as well as respiratory rate varies with SL when repeatedly anaesthetised using protocol 1 every 4 days. Scatter plot of (A) induction time, (B) time taken to lose touch responsivity, (C) beats per minute at 5:00, (D) breaths per minute at 5:00, (E) time taken to regain movement and (F) time taken to fully recover against repeat number over their standard length (mm). Blue dots correspond to fish that were repeatedly dosed. Red dots correspond to fish that were dosed once (*i.e.* control group). Blue and red linear regression lines are superimposed with translucent 95% confidence interval.



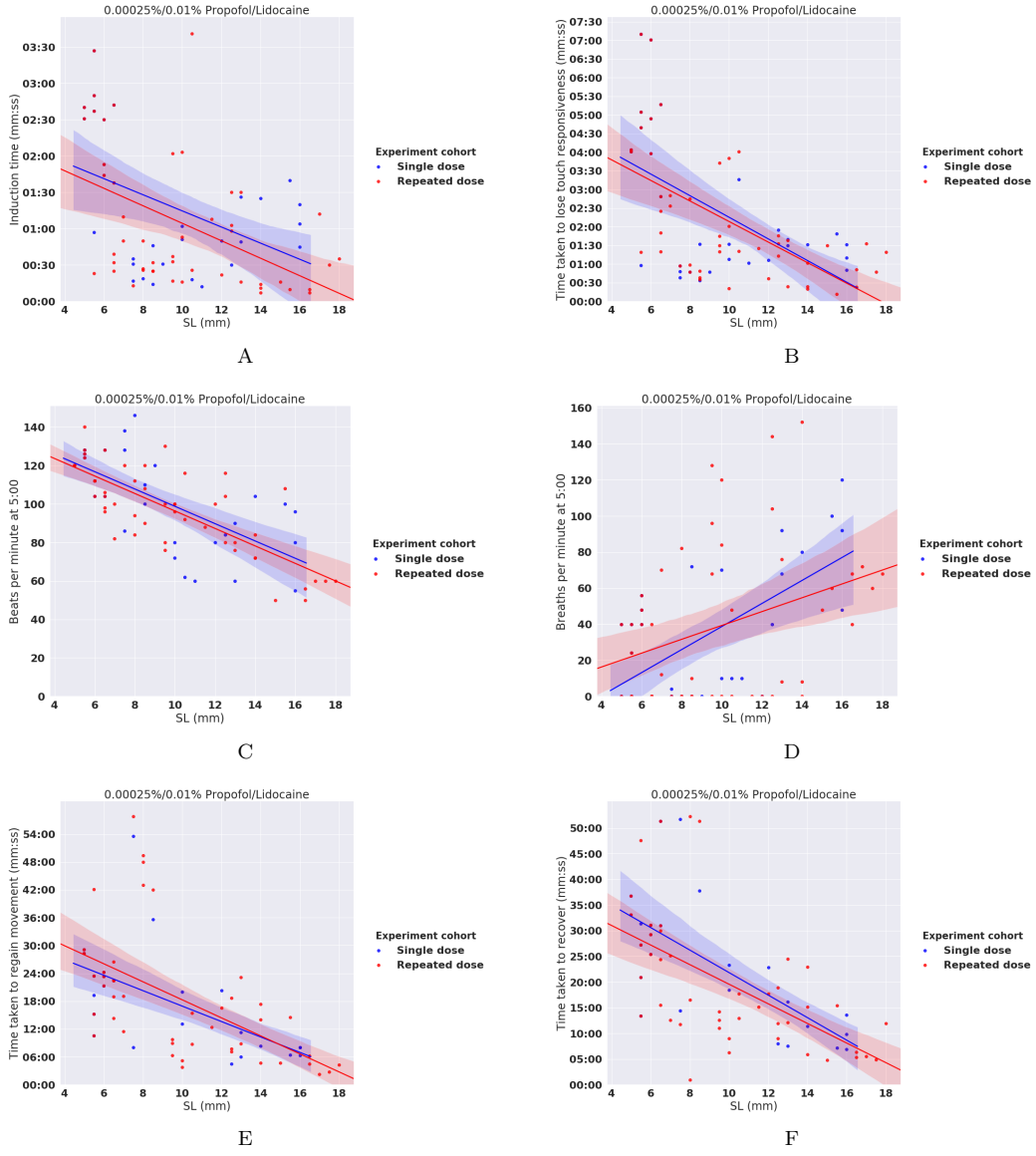
Supplementary Fig. 4: The time taken to induce, lose touch responsiveness, regain movement, recover as well as respiratory rate varies with repeat number when repeatedly anaesthetised using protocol 4 every 4 days. Boxplot of (A) induction time, (B) time taken to lose touch responsivity, (C) heartbeats per minute at 5:00, (D) visible breaths per minute at 5:00, (D) time taken to regain movement and (E) time taken to fully recover against repeat number when using Protocol 4 to repeatedly anaesthetise and recover fish. Stars indicate significant difference as determined using the Mann-Whitney U test. ns = not significant, '*' indicates $p<0.05$, '**' indicates $p<0.01$, '***' indicates $p < 0.001$ and '****' indicates $p<0.0001$. CP stands for cotton pad maintenance method.



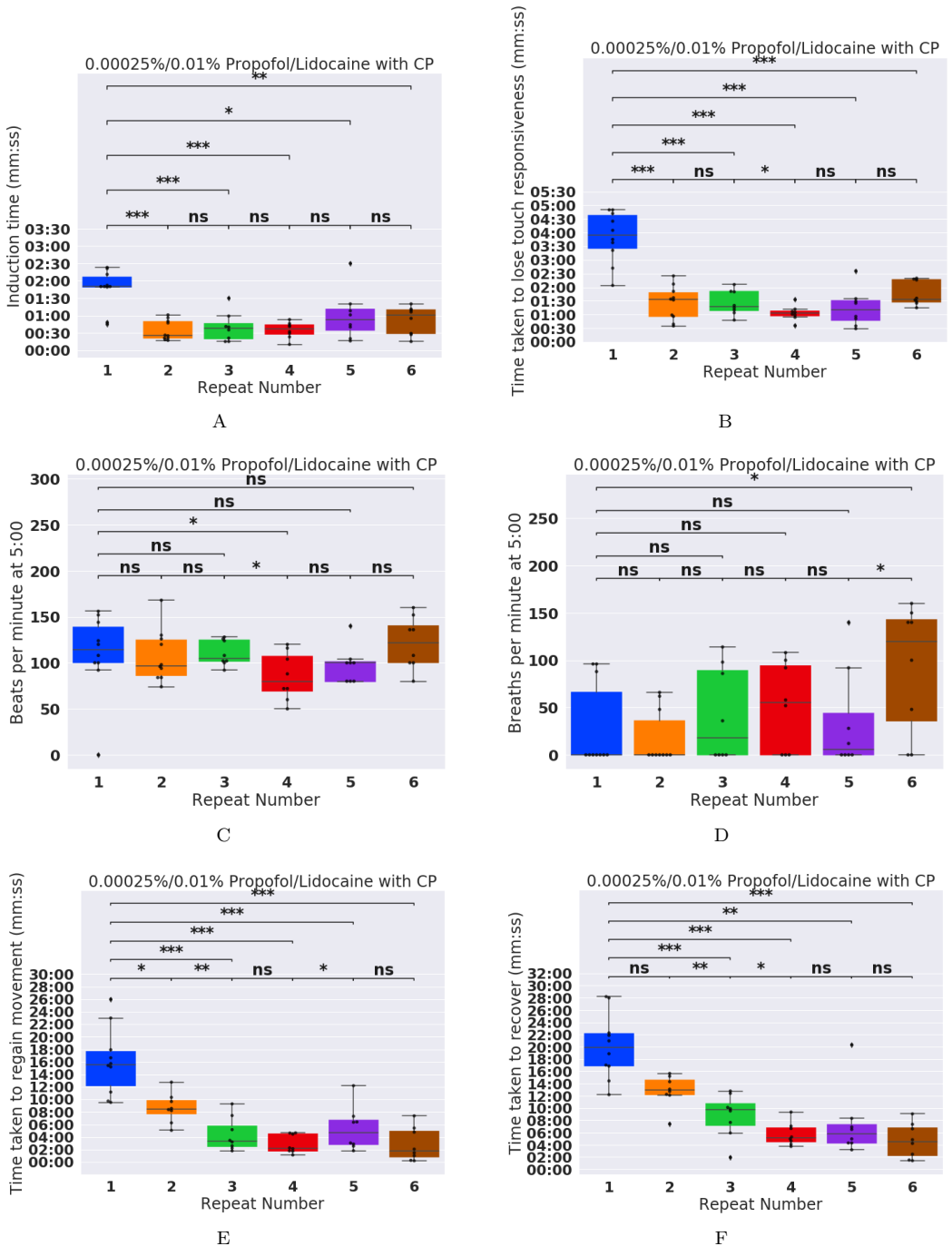
Supplementary Fig. 5: The time taken to induce, lose touch responsiveness, regain movement, recover as well as respiratory rate varies with SL when repeatedly anaesthetised using protocol 4 every 4 days. Scatter plot of (A) induction time, (B) time taken to lose touch responsivity, (C) beats per minute at 5:00, (D) breaths per minute at 5:00, (E) time taken to regain movement and (F) time taken to fully recover against repeat number over their standard length (mm). Blue dots correspond to fish that were repeatedly dosed. Red dots correspond to fish that were dosed once (*i.e.* control group). Blue and red linear regression lines are superimposed with translucent 95% confidence interval. CP stands for cotton pad maintenance method.



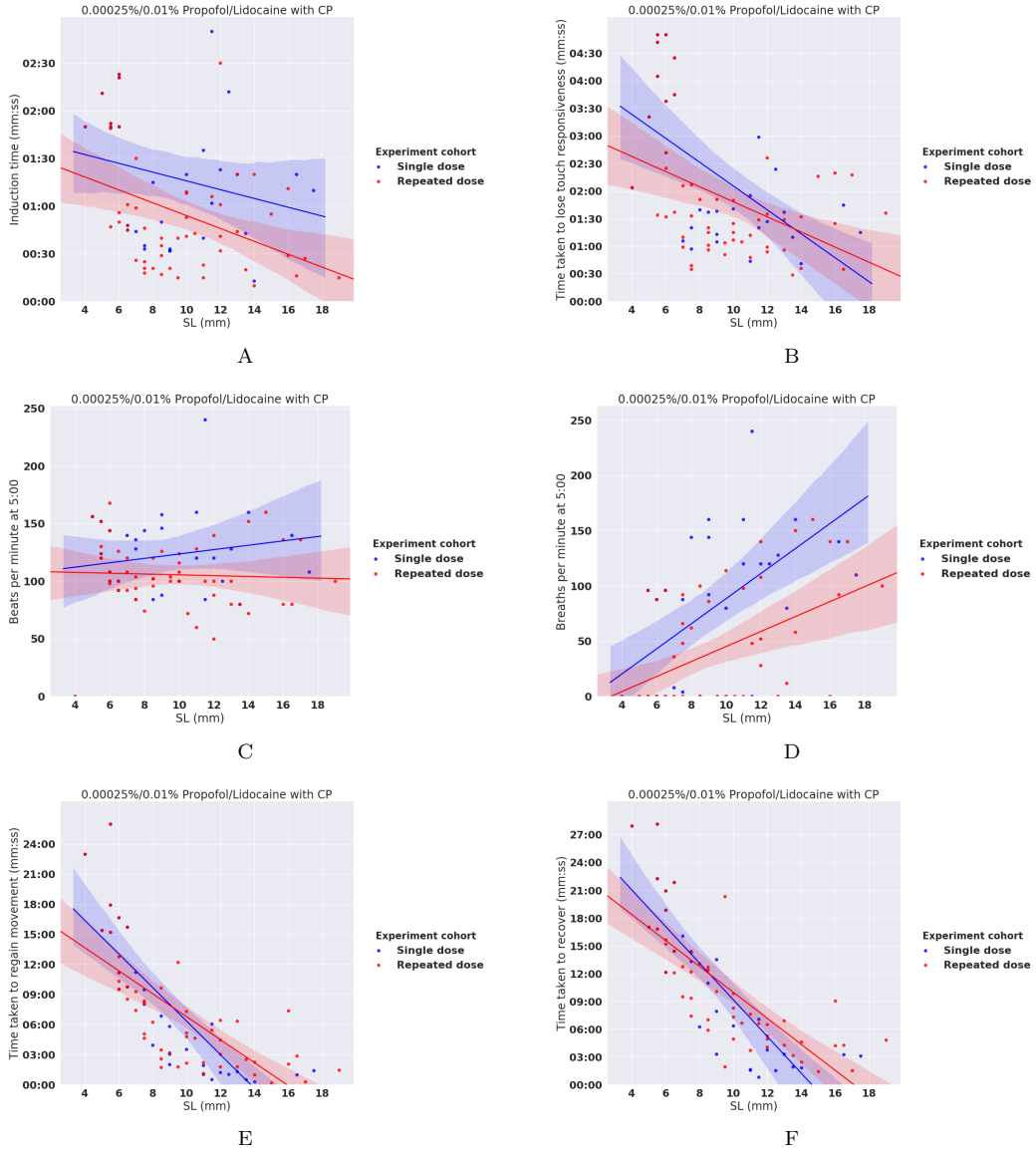
Supplementary Fig. 6: The time taken to induce, lose touch responsiveness, regain movement, recover as well as respiratory rate varies with repeat number when repeatedly anaesthetised using protocol 5 every 4 days. Boxplot of (A) induction time, (B) time taken to lose touch responsivity, (C) heartbeats per minute at 5:00, (D) visible breaths per minute at 5:00, (D) time taken to regain movement and (E) time taken to fully recover against repeat number when using Protocol 5 to repeatedly anaesthetise and recover fish. Stars indicate significant difference as determined using the Mann-Whitney U test. ns = not significant, '*' indicates $p<0.05$, '**' indicates $p<0.01$, '***' indicates $p < 0.001$ and '****' indicates $p<0.0001$.



Supplementary Fig. 7: The time taken to induce, lose touch responsiveness, regain movement, recover as well as respiratory rate varies with SL when repeatedly anaesthetised using protocol 5 every 4 days. Scatter plot of (A) induction time, (B) time taken to lose touch responsivity, (C) beats per minute at 5:00, (D) breaths per minute at 5:00, (E) time taken to regain movement and (F) time taken to fully recover against repeat number over their standard length (mm). Blue dots correspond to fish that were repeatedly dosed. Red dots correspond to fish that were dosed once (*i.e.* control group). Blue and red linear regression lines are superimposed with translucent 95% confidence interval.



Supplementary Fig. 8: The time taken to induce, lose touch responsiveness, regain movement, recover as well as respiratory rate varies with repeat number when repeatedly anaesthetised using protocol 6 every 4 days. Boxplot of (A) induction time, (B) time taken to lose touch responsivity, (C) heartbeats per minute at 5:00, (D) visible breaths per minute at 5:00, (D) time taken to regain movement and (E) time taken to fully recover against repeat number when using Protocol 6 to repeatedly anaesthetise and recover fish. Stars indicate significant difference as determined using the Mann-Whitney U test. ns = not significant, '*' indicates $p<0.05$, '**' indicates $p<0.01$, '***' indicates $p < 0.001$ and '****' indicates $p<0.0001$. CP stands for cotton pad maintenance method.



Supplementary Fig. 9: The time taken to induce, lose touch responsiveness, regain movement, recover as well as respiratory rate varies with SL when repeatedly anaesthetised using protocol 6 every 4 days. Scatter plot of (A) induction time, (B) time taken to lose touch responsivity, (C) beats per minute at 5:00, (D) breaths per minute at 5:00, (E) time taken to regain movement and (F) time taken to fully recover against repeat number over their standard length (mm). Blue dots correspond to fish that were repeatedly dosed. Red dots correspond to fish that were dosed once (*i.e.* control group). Blue and red linear regression lines are superimposed with translucent 95% confidence interval. CP stands for cotton pad maintenance method.

Supplementary Material (Tables)

S1 Table: The effects of repeat number and SL on success measures for protocol 1 when used every 4 days. Multiple linear regression model for y= induction time, time taken to lose touch responsiveness, beats per minute, breaths per minute, time taken to regain movement and time taken to recover under protocol 1. For each y value a significant regression equation was found if $p<0.05$ and can be described as (F(Df model, Df residual)=F-stat. The predicted y value is given by $y=C+m_1*(\text{repeat number})+m_2*(\text{standard length})$, where repeat number is the number of doses the fish will have been exposed to at the end of the experiment and standard length is in mm. Values highlighted are those where the associated p-value is <0.05 and thus can be deemed as significant.

Protocol 1	**Df res.**	**Df Model**	**F-stat**	**p**	**R^2**	**C**	**Repeat number coef. (m_1)**	**Repeat number p-value**	**SL coef. (m_2)**	**SL p-value**
Induction time (s)	87	2	73.20	<0.05	0.63	-53.46	7.39	<0.05	13.32	<0.05
Time taken to lose touch responsiveness (s)	87	2	64.85	<0.05	0.60	-13.94	7.91	<0.05	11.09	<0.05
Beats per minute at 5:00	87	2	10.36	<0.05	0.19	91.77	-3.67	0.16	6.73	<0.05
Breaths per minute at 5:00	87	2	22.29	<0.05	0.34	23.01	-2.26	0.53	12.22	<0.05
Time taken to regain movement (s)	87	2	11.95	<0.05	0.22	110.81	-8.45	<0.05	-2.04	0.24
Time taken to recover (s)	87	2	14.88	<0.05	0.26	132.09	-9.78	<0.05	-2.35	0.19

S2 Table: The effects of repeat number and SL on success measures for protocol 4 when used every 4 days. Multiple linear regression model for y= induction time, time taken to lose touch responsiveness, beats per minute, breaths per minute, time taken to regain movement and time taken to recover under protocol 1. For each y value a significant regression equation was found if $p<0.05$ and can be described as (F(Df model, Df residual)=F-stat. The predicted y value is given by $y=C+m_1*(\text{repeat number})+m_2*(\text{standard length})$, where repeat number is the number of doses the fish will have been exposed to at the end of the experiment and standard length is in mm. Values highlighted are those where the associated p-value is <0.05 and thus can be deemed as significant.

Protocol 4	**Df res.**	**Df Model**	**F-stat**	**p**	**R^2**	**C**	**Repeat number coef. (m_1)**	**Repeat number p-value**	**SL coef. (m_2)**	**SL p-value**
Induction time (s)	87	2	7.36	<0.05	0.15	22.61	-1.65	<0.05	-0.08	0.82
Time taken to lose touch responsiveness (s)	87	2	19.02	<0.05	0.30	38.37	-1.40	<0.05	-1.00	<0.05
Beats per minute at 5:00	87	2	4.48	<0.05	0.09	22.83	2.77	0.43	3.80	0.08
Breaths per minute at 5:00	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Time taken to regain movement (s)	87	2	13.37	<0.05	0.24	26.31	-27.76	<0.05	27.43	<0.05
Time taken to recover (s)	87	2	12.94	<0.05	0.23	30.86	-24.99	<0.05	29.85	<0.05

S3 Table: The effects of repeat number and SL on success measures for protocol 5 when used every 4 days. Multiple linear regression model for y= induction time, time taken to lose touch responsiveness, beats per minute, breaths per minute, time taken to regain movement and time taken to recover under protocol 1. For each y value a significant regression equation was found if $p<0.05$ and can be described as (F(Df model, Df residual)=F-stat. The predicted y value is given by $y=C+m_1*(\text{repeat number})+m_2*(\text{standard length})$, where repeat number is the number of doses the fish will have been exposed to at the end of the experiment and standard length is in mm. Values highlighted are those where the associated p-value is <0.05 and thus can be deemed as significant.

Protocol 5	**Df res.**	**Df Model**	**F-stat**	**p**	**R^2**	**C**	**Repeat number coef. (m_1)**	**Repeat number p-value**	**SL coef. (m_2)**	**SL p-value**
Induction time (s)	77	2	11.86	<0.05	0.236	136.67	-6.02	0.16	-5.38	<0.05
Time taken to lose touch responsiveness (s)	77	2	20.13	<0.05	0.343	297.10	-7.74	0.32	-14.71	<0.05
Beats per minute at 5:00	77	2	42.48	<0.05	0.525	142.27	-1.21	0.39	-4.20	<0.05
Breaths per minute at 5:00	77	2	9.89	<0.05	0.204	-10.96	-3.40	0.28	5.79	<0.05
Time taken to regain movement (s)	67	2	15.60	<0.05	0.318	2173.49	28.38	0.63	-117.71	<0.05
Time taken to recover (s)	67	2	20.52	<0.05	0.380	2445.24	-2.69	0.96	-121.37	<0.05

S4 Table: The effects of repeat number and SL on success measures for protocol 6 when used every 4 days. Multiple linear regression model for y= induction time, time taken to lose touch responsiveness, beats per minute, breaths per minute, time taken to regain movement and time taken to recover under protocol 1. For each y value a significant regression equation was found if $p<0.05$ and can be described as (F(Df model, Df residual)=F-stat. The predicted y value is given by $y=C+m_1*(\text{repeat number})+m_2*(\text{standard length})$, where repeat number is the number of doses the fish will have been exposed to at the end of the experiment and standard length is in mm. Values highlighted are those where the associated p-value is <0.05 and thus can be deemed as significant.

Protocol 6	**Df res.**	**Df Model**	**F-stat**	**p**	**R^2**	**C**	**Repeat number coef. (m_1)**	**Repeat number p-value**	**SL coef. (m_2)**	**SL p-value**
Induction time (s)	79	2	6.69	<0.05	0.15	97.96	-4.97	0.09	-2.31	0.11
Time taken to lose touch responsiveness (s)	79	2	13.68	<0.05	0.26	213.32	-7.36	0.14	-7.88	<0.05
Beats per minute at 5:00	79	2	1.71	0.19	0.04	107.37	-4.79	0.07	1.71	0.20
Breaths per minute at 5:00	79	2	20.15	<0.05	0.34	-24.41	-14.05	<0.05	12.18	<0.05
Time taken to regain movement (s)	77	2	48.89	<0.05	0.56	1203.15	5.42	0.78	-81.56	<0.05
Time taken to recover (s)	67	2	20.52	<0.05	0.62	2445.24	-2.69	0.96	-121.37	<0.05

S5 Table: The effects of repeat number and SL on success measures for protocol 4 when used every 2 days. Multiple linear regression model for y= induction time, time taken to lose touch responsiveness, beats per minute, breaths per minute, time taken to regain movement and time taken to recover under protocol 4 when used every 2 days. For each y value a significant regression equation was found if $p<0.05$ and can be described as (F(Df model, Df residual)=F-stat. The predicted y value is given by $y=C+m_1*(\text{repeat number})+m_2*(\text{standard length})$, where repeat number is the number of doses the fish will have been exposed to at the end of the experiment and standard length is in mm. Values highlighted are those where the associated p-value is <0.05 and thus can be deemed as significant.

Protocol 4 (Every 2 days)	**Df res.**	**Df Model**	**F-stat**	**p**	**R^2**	**C**	**Repeat number coef. (m_1)**	**Repeat number p-value**	**SL coef. (m_2)**	**SL p-value**
Induction time (s)	87	2	7.89	<0.001	0.154	14.4	-4.666	<0.05	2.69	<0.05
Time taken to lose touch responsiveness (s)	87	2	16.26	<0.001	0.272	29.26	-6.25	<0.05	2.96	<0.05
Time taken to regain movement (s)	87	2	17.24	<0.001	0.284	211.3	-8.48	0.07	-4.61	0.18
Time taken to recover (s)	87	2	14.39	<0.001	0.249	232.4	-12.86	0.02	-1.98	0.62

7.11. Conclusions

In this chapter we presented our work determining a suitable anaesthetic protocol for repeatedly anaesthetising and recovering juvenile zebrafish every 2 days. The aim of this paper was to refine the method used to anaesthetise juvenile zebrafish, so that the fish could be repeatedly imaged with minimal suffering.

The protocol determined in this chapter should be considered a refinement to the previously adult-focussed methods for anaesthetising juvenile zebrafish. The results of this paper will not only aid the documentation of adult stripe pattern formation [46, 31] but also the many other developmental milestones the juvenile stages encompasses, such as gill ontogenesis [129], skin maturation including scale growth and skin innervation [83]. Furthermore, the results will help to minimise suffering and thus contributes to the goals of National Centre for the Replacement, Refinement and Reduction of Animals in Research. Furthermore, by drawing attention to the potentially harmful effects of using common fish anaesthetic: MS-222 during this sensitive stage, we will reduce the number of zebrafish affected negatively by using MS-222 naively.

The results of this paper build upon previous work by testing different methods of anaesthetising fish that had been deemed suitable for adult fish during this sensitive juvenile period [123, 122]. We demonstrate that most of the methods suitable for adult fish can be adapted for juvenile fish by changing the dose and/ or method. It is also the first paper to investigate the effects of repeated anaesthesia on zebrafish during development. We demonstrate that whilst repeated anaesthesia using any of the four protocols that we tested did not stunt the development of the fish that, repeated anaesthesia could have a significant effect on other aspects of the protocol, such as the time taken to anaesthetise and recover the fish.

In the context of this thesis, the results of this paper are inspired by and applicable to Chapter 8. In Chapter 8 we use the protocol in order to repeatedly anaesthetise and image fish aged between 23 dpf and 40 dpf. From the images that we collate, we are able to determine the relative position of the X0 over time, parameterising my model accordingly and subsequently aiding our investigation into biased growth during zebrafish pattern formation. Extensive studies of the same nature, could be used to parameterise many other aspects of pattern development, such as the speed of iridophore movement, the rate of death of various cell types as well as the pattern development of some undocumented mutants, to name just a few.

Chapter 8

Differential growth is a critical determinant of zebrafish pigment pattern formation

8.1. Overview

This chapter builds upon previous work, using the model developed in Chapter 6 along with experimental results generated using techniques defined in Chapter 7 to investigate the importance of growth to pattern development in zebrafish. The chapter comprises a paper written by myself (Jennifer Owen), Christian Yates and Robert Kelsh that is yet to be submitted. In this chapter, we use the mathematical model specified in Chapter 6 in order to investigate the impact of growth on zebrafish stripe formation.

The motivation for this chapter comes from observations of adult zebrafish, which despite displaying an initially symmetric pattern in the dorso-ventral (DV) axis, are characteristically asymmetric in adulthood. For example, the X0 interstripe which is initially centred on the domain, moves dorsally relative to the full height of the fish by the time of adulthood. Mutants displaying spots on the body generate spots that are more rounded ventrally and more elongated and stripe-like dorsally. Since the cell-cell interactions involved in pattern development are likely symmetric (see Chapter 3), we predict that this effect is caused by growth. More specifically we predict that growth in *Danio rerio* is biased to occur more frequently in the ventral region.

In the first part of this paper, building upon observations published in eLife (see Chapter 6) we demonstrate the importance of growth to the pattern. By changing the strength of the growth in different directions we show that we can generate many different patterns including those of *Danio* relatives. In the latter part of the paper we perform an interdisciplinary study into the asymmetry of zebrafish stripe patterns. Motivated by our prediction that this is caused by ventral growth bias, we define a one-dimensional model of growth in the dorso-ventral direction which we use to parameterise growth bias in our two-dimensional model. We show that by incorporating biased growth, not only do our simulated patterns of WT and missing-cell mutants more qualitatively and quantitively resemble real fish, but we are able to reproduce mutants and relatives of *Danio rerio* that had not been possible before, such as alleles of *rose* and *Danio nigrofasciatus*.

8.2. Outline of the chapter

Section I of the paper provides a introduction to the importance of studying growth in the context of understanding pattern formation. In Section II we outline the key results of this study: namely that by incorporating growth that is biased to occur more frequently in the ventral region into our model of zebrafish stripe formation that we can explain many interesting facets of zebrafish pattern development, including the asymmetry of zebrafish patterns as well other quantitative details such as melanophore position change during development. We discuss the importance of our results in context in Section III. In Section IV we provide a brief outline of our mathematical model as part of the materials and methods. The supplementary material contains the derivations of the one-dimensional model that we use to parameterise our comprehensive

model but are omitted from the main text.

Differential growth is a critical determinant of zebrafish pigment pattern formation

Jennifer Owen[1] , Robert N Kelsh[1]* and Christian Yates[2]*
1 Department of Biology and Biochemistry, University of Bath, Bath, United Kingdom
2 Department of Mathematical Sciences, University of Bath, Bath, United Kingdom

8.3. Summary paragraph

The coloured skin patterns of vertebrates are formed by complex interactions between pigment-producing cells during development. The adult zebrafish (*Danio rerio*), a model organism for investigating the underlying patterning processes, possesses alternating horizontal blue and golden stripes, generated by the self-organisation of three pigment cell-types over three weeks during development. Mathematical studies in which the three populations are modelled as individual agents communicating via short- and long-range interactions have produced breakthroughs in the understanding of striped pattern development in zebrafish. These models, which incorporate all experimentally evidenced cell-cell interactions, replicate many aspects of wild-type and mutant zebrafish patterns. Although received wisdom suggested that initial iridophore stripe formation was pivotal in orienting patterning, we show that growth can override the influence of such initial conditions, with altered growth sequences generating further pattern modulation, including vertical stripes and maze-like patterns. Here we test the hypothesis that asymmetric growth of the skin field might elucidate two key zebrafish pattern features which are otherwise obscure: dorso-ventral pattern asymmetry, and the predominant ventral-to-dorsal migration of melanophores during pattern development. Our findings suggest that ventrally biased (*i.e.* asymmetric) growth can indeed explain these features in wild-type and multiple zebrafish mutants, and in the related species *Danio nigrofasciatus*. By identifying biased growth as a novel patterning mechanism, our study will inform future investigations into the mechanisms and evolution of fish pigment patterning and the general development of vertebrate pigment pattern formation. Furthermore, our work has implications for the mechanistic basis of human pigmentation defects.

8.4. Introduction

The diversity of skin pigment patterns in zebrafish (including those of numerous pigment pattern mutants) and related Danio species, combined with a wealth of experimental observations, make the zebrafish a paradigmatic example of pigment pattern formation. Zebrafish pigment pattern formation involves diverse cell-cell interactions, which generate both wild-type and mutant pigment patterns [46, 145, 64, 146, 41, 39, 43, 26, 75, 147]. Theoretical modelling strongly suggest that this occurs via a complex, but defined, network of interactions [28, 26], whilst experimental studies indicate that the initial horizontal orientation along the mid-flank of the first iridophores determines the orientation of the horizontally-striped pattern [46, 18, 39]. However, our initial modelling indicated this might not always be sufficient [28] and the marked dorso-ventral (DV) asymmetries in zebrafish pigment patterns remain unexplained. During pattern development, zebrafish grow up to 7.5 fold in length and 200 fold in volume,, at a rate highly dependent on rearing conditions but at least at $\approx$130 μm per day in the anteroposterior direction and $\approx$33 μm per day in the DV direction [31]. The growth rate influences the success of stripe formation in some mutants [27]. There is some correlation between Danio species pigment and growth patterns with, for example, species

with vertical disruptions in the stripes having deeper bodies and a prolonged period of juvenile growth [148]. Building on our model of zebrafish pattern development [28], we show that growth rates are key determinants of patterning, readily over-riding initial iridophore patterns. By changing growth parameters, whilst keeping pigment cell interactions and initial conditions unchanged, we were able to generate diverse patterns resembling those of other species within the *Danio* genus. Furthermore, altering the degree of asymmetry of growth allowed us to reproduce the DV asymmetry of *Danio* patterns, leading us to predict that DV growth in Danio rerio is biased towards ventral regions, and may also explain the previously unexplored DV asymmetry of *Danio nigrofasciatus*.

8.5. Results

8.5.1 Initial iridophore position is not predictive of the orientation of pattern

To test the hypothesis that the horizontal orientation of the initial iridophore stripe determines the longitudinal orientation of the final pigment pattern in zebrafish, we performed an in silico experiment using our agent-based model to simulate development up to the J+ stage [28, 31]. As demonstrated previously [28], with the initial condition set to include 3 rows of iridophores horizontally across the center of the domain as seen in WT fish, we consistently observe horizontally-organised stripes at the end of simulations with WT growth parameters (Fig. 8-1A and Fig. 8-1A', Supplementary Movie 1). Surprisingly, even when the initial position of the iridophores in our simulation was dramatically altered (whilst all other conditions remained unchanged), the simulated patterns consistently showed distinct horizontal stripe elements in the final orientation of the pattern 30 days later (Fig. 8-1B-1D, Fig. 8-1B'-D' and Supplementary Movies 2-4). For example, a vertical initial iridophore stripe does not force a vertical barred pattern, but instead generates a broken pattern, but with strong horizontal elements (Fig. 8-1B); a cross-shaped combination of horizontal and vertical initial iridophores results in largely normal pattern, albeit with central breaks to the most dorsal and ventral stripes (Fig. 8-1C), whilst a central square of initial iridophores produces an intermediate pattern with both horizontal and vertical stripe elements (Fig. 8-1D). To provide an objective quantification of striping, we used a pair correlation function (PCF) to determine and quantify spatial correlations by taking the mean of the final outputs of the PCF [103] in the horizontal (and separately vertical) direction for 50 repeats, for each of the initial conditions. Regardless of the initial iridophore distribution, there was strong pairwise correlation in the vertical direction, but not the horizontal (Fig. 8-1A, Fig. 8-1B, Fig. 8-1C and Fig. 8-1D), showing that the orientation of the patterns formed at stage J+ were all dominated by horizontal elements. These results directly contradict the hypothesis that the initial interstripe orientation plays the defining role in driving pattern orientation.

8.5.2 Changing the rate of growth markedly changes patterning

We observed that simulated vertical pattern elements appeared consistently to correlate with domain growth (Supplementary Movie 2). To test whether this correlative relationship was causal, we performed simulations where the initial iridophore interstripe was maintained horizontally, but the growth rates in horizontal and vertical axes were changed (Supplementary Movies 5-9). First, we set the growth rate to be equal

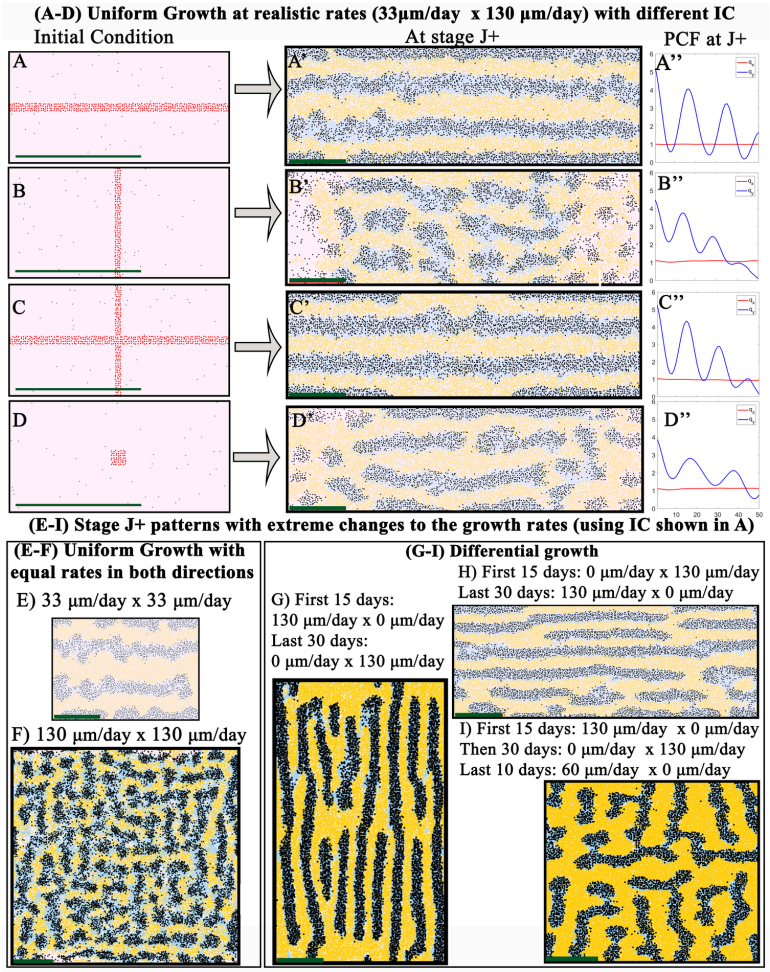


Figure 8-1: ***Pigment pattern orientation depends on growth, not just orientation of initial interstripe.*** *(A) - (D) Hypothetical initial conditions for the position of dense iridophores (coloured in red for clarity). (A) WT Danio rerio initial conditions (B) Vertical stripe of iridophores (C) Cross shape of iridophores. (D). Square of iridophores in the center (A') - (D') The predicted pattern at stage J+ given respective initial conditions and normal growth rates: 33µm/day in the vertical direction and 130µm/day in the horizontal direction. (A") - (D") The averaged PCF of 50 repeats of stage J+ simulations with IC given in (A) - (D) in the horizontal axis (red) and vertical axis (blue) using the PCF of Binder et al 2013 [105]. (E) (I) The predicted patterns at stage J+ with WT Danio rerio initial conditions shown in (A) but different rates of growth: (E) 33µm/day in the horizontal axis and y direction. (F) 130um/day in the horizontal and vertical direction. (G) 130µm/day in the vertical direction for 15 simulated days (and 0µm/day in the horizontal direction) followed by 130µm/day in the horizontal direction for the following 30 simulated days (and 0µm/day in the vertical direction). (H) The predicted pattern at stage J+ when growth is 130µm/day in the horizontal direction (and 0µm/day in the vertical direction) for 15 simulated days followed by 130µm/day in the vertical direction (and in the horizontal direction) for the following 30 simulated days. (I) the predicted pattern at stage J+ when growth is 130µm/day in the horizontal direction (and 0µm/day in the vertical direction) for 15 simulated days followed by 130µm/day in the vertical direction (and 0µm/day in the horizontal direction) for the following 30 simulated days, followed by 60 µm/day in horizontal direction (and 0µm/day in the vertical direction) for 10 days. The green scale bar indicates 1mm in all figures. Alternative simulations for the simulations represented in (A)-(I) are given in Supplementary Movies 1-9.*

in both directions; when the growth rate was slow stripes were maintained (Fig. 8-1E; Supplementary Movie 5), but when growth was faster, the pattern degenerated into a broken labyrinthine configuration with a tendency towards spots (Fig. 8-1F; Supplementary Movie 6). We then explored the impact of time-dependent growth rates. Setting growth rates to 130μm per day in the vertical direction (and 0μm per day in the horizontal direction) for the first 15 days and then 130μm per day in the horizontal direction (and 0μm per day in the vertical direction) for the latter 30 days) generated a pattern of fine vertical bars (Fig. 8-1G; Supplementary Movie 7); examination of the time-course showed a clear switch from early horizontal striping to increasingly vertical barring triggered by the change in growth orientation (Supplementary Movie 7). In contrast, reversing this growth pattern (130μm per day in the horizontal direction for 15 simulated days (and 0μm per day in the vertical direction), and then switching to 130μm per day in the vertical direction for 30 simulated days (and 0μm per day in the horizontal direction)), generated a maze-like pattern (Fig. 8-1H; Supplementary Movie 8): stripes broken into spots with smooth edges; examination of the time course reveals transient horizontal stripes being overwhelmed by the vertical growth to become rapidly a vertically-barred pattern, before subsequently being modified to form extensive horizontal stripe elements upon the switch to horizontal growth (Supplementary Movie 8). When growth alternated between being exclusively in the horizontal and vertical directions multiple times, simulations generated a labyrinthine pattern with strong horizontal and vertical stripe elements, again driven by the changing growth orientations (Fig. 8-1I; Supplementary Movie 9). Thus, growth rates are likely to play a key role in controlling major pattern features.

8.5.3 Impact of biased growth on multiple features of zebrafish pigment pattern

We hypothesised that asymmetric growth in the DV axis might explain the apparent dorsal shift of the X0 interstripe during metamorphosis [31], and the relative timing and relative width of secondary stripe formation. To assess asymmetry in body growth during pigment pattern formation, we repeatedly measured the dorso-ventral position (DVP) of the centre of the X0 interstripe of 8 individual zebrafish during the developmental period 8 mm SL 16 mm SL. We found that the DVP of the centre of X0 moves from relative position 0.5±0.01 at stage PB (8mm SL), to relative position 0.4±0.01 by stage J+ (16 mm SL,shown quantitatively in Fig. 8-2A; rate of translocation quantified in Fig. 8-2B). These observations suggest that growth might be ventrally-biased in the DV axis. Further to this, during pattern metamorphosis, stripes and interstripes form sequentially from the central interstripe. Assuming symmetric growth about the DV axis (and all cell-cell interactions are symmetric), we would expect that stripes 2D and 2V would appear at the same time. We evaluated the time at which stripes 1D, 1V, 2D and 2V first become visible on the fish with respect to developmental stages corresponding to size. We found that whilst in all cases stripes 1D and 1V both became visible simultaneously (mean±sd: 9.69 ± 0.45 mm SL), stripe 2V always appeared before stripe 2D (2V; 11.81±0.84 mm SL, 2D; 14.69±1.16 mm SL, Fig. 8-2C). This difference is most readily observed when assessing development of the stripes close to the caudal fin where the fish is narrower (Fig. 8-2D). At 10 mm SL, the domain comprises the two stripes 1D and 1V at the very dorsal and ventral regions, respectively, separated by the X0 interstripe. As the fish grows bigger, whilst stripe 1D remains at the very dorsal region of the fish, stripe 1V relocates dorsally over time. By 16 mm SL the 1V stripe is close to being aligned with the central caudal fin stripe, allowing

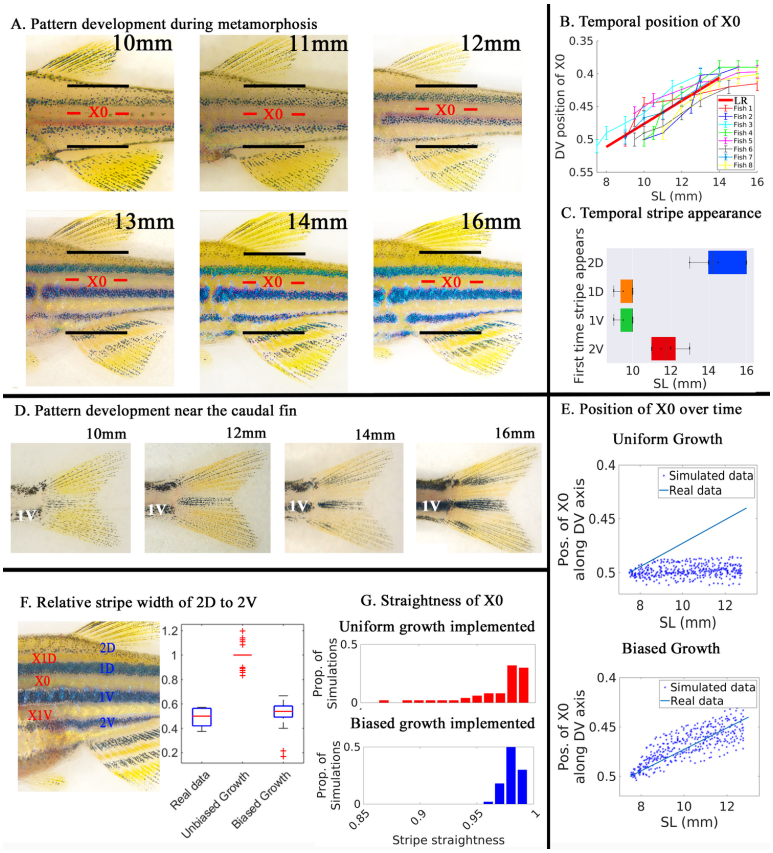


Figure 8-2: ***Dorsal repositioning of the first (X0) interstripe can be explained by ventrally-biased growth during pattern development.*** *(A) WT pattern development on the body. The black lines indicate the dorsal and ventral extremes of posterior trunk. The red line indicates the DV midline of the X0 interstripe. Note how the X0 interstripe, which is initially central within the DV axis, translocates dorsally over time. (B) Shown is the relative position of X0 along the DV axis against the standard length for each fish ($N = 8$). Thin coloured lines plot measured values of each fish The thick red line represents the results of undertaking a linear regression (LR) of all of the points as described in the Materials and Methods. (C) Boxplot showing the temporal appearance of stripes during development. Lines indicate in order the minimum value, lower quartile, median, upper quartile and finally maximum value. Shown is the length of the fish on the first day of imaging that each stripe was first visible. Stripes 1D and 1V always first became visible on the same day. Stripe 2D always became visible a few days later than stripe 2V. (D) Pattern development on the body close to the caudal fin. Stripe 1V marked by the label, is initially adjacent to the most ventral part of the fish at 10 mm SL but has shifted to a central position by 16 mm SL; note the shift in alignment to the central stripe on the caudal fin itself. (E) Comparison of the variation of the simulated position of X0 with standard length using simulations of WT fish with unbiased (upper) and biased growth (lower). The blue line indicates the line of best fit for the real data and is the same as the red line shown in (B). Simulations with ventrally-biased growth show a significantly better match to the real data than those with unbiased growth. (F) Juvenile fish at stage J+ (13 mm SL) wherein stripes and interstripes have been labelled accordingly; stripe 2D is consistently thinner than stripe 2V. Graph shows comparison of the ratio of DV extent of stripes 2D to 2V with real data and 50 simulations of WT fish at stage J+ using biased and unbiased growth parameters. A box plot of the ratio of 2D to 2V with real data and 50 simulations of WT fish at stage J+ using biased and unbiased growth parameters. There is a significantly better match to the real data using biased growth, than using unbiased growth. (G) A comparison of the straightness of X0 of WT simulations at stage J+ using biased and unbiased growth parameters. WT simulations with biased growth show an increase in the mean stripe straightness as well as a decrease in the variance in the stripe straightness compared with simulations using unbiased growth.*

space for a new 2V stripe to appear ventrally. Meanwhile 1D remains at the very dorsal part of the domain. Implementing our model (Owen et al., 2020) with unbiased growth (Fig. 8-2E, upper panel) generates a poor match of the position of X0 during later growth stages in the model compared to the measured X0 position; indeed, the modelled position of X0 remains constant in the middle of the domain (relative dorso-ventral position 0.5). We adapt the model so that the domain grows more ventrally than dorsally (see Supplementary Material). Using the equation relating the position of X0 and the standard length measured from eight fish (derived using linear regression), we determine parameters for biased growth that give a good fit to the real data based on the position of X0 over time (see Supplementary Material). The impact of this bias on X0 position is displayed in Fig. 8-2E (lower panel). The position of X0 migrates dorsally in our simulations as the fish grows larger, at a rate similar to that quantified in real fish.

One further significant limitation of our previous unbiased simulations was the failure to predict the dorsoventral asymmetries in stripe width [28]. We now show that this is corrected for by implementing ventrally biased growth. In real fish, while stripes 1D and 1V are of approximately equal width, stripe 2D is 0.52±0.05 times as thick as stripe 2V. We propose that this results from biased growth, which increases the space available for stripe 2V to develop relative to stripe 2D. In Fig. 8-2F we compare the ratio of the width of 2D to the width of 2V in real fish, with those seen in simulations employing either unbiased or biased growth. Using unbiased growth, the mean stripe width ratio of 2D and 2V is, unsurprisingly, 1±0.02, which does not match the real data. In contrast, implementing biased growth the mean stripe width ratio is 0.58±0.05, similar to the real value. Interestingly, we found that biasing the growth ventrally also increased the mean straightness of the X0 interstripe, as well as reducing the variance (biased growth: 98±0.08%, unbiased growth: 96.9±2.6%)

Finally we present ventrally biased growth as an alternative explanation for observed changes in melanophore differentiation and location during metamorphosis. Parichy and Turner document the initial relative dorso-ventral position (RDVP) of melanophores observed over the course of development (Fig. 8-3A) [27]. They showed that melanophore births which are initially uniformly distributed between 0-0.25 RDVP and 0.75-1 RDVP, become biased ventrally during pattern metamorphosis. We predicted that this observation might result from ventrally biased growth. Consistent with this, the position of melanophore differentiation in simulations using our model with unbiased growth do not match the published observations; instead, melanophore differentiation is symmetric along the DV axis (Fig. 8-3A). In contrast, when we incorporate biased growth (Fig. 8-3A)) we observe the same ventral bias as documented in real fish (Fig. 8-3A), reflecting the dorsal shift of the X0 stripe.

Parichy and Turner also document the net change in RDVP of melanophores from their initial RDVPs at the time of initial differentiation [27] (Fig. 8-3B), observing that melanophores tended to relocate dorsally during pattern development and that this was most prominent for the melanophores comprising stripe 1V. They interpreted this change as resulting from dorsally-biased migration. Whilst this could undoubtedly be achieved through a complicated mechanism specifying biased migration of certain cells in certain parts of the domain during certain time periods, a more parsimonious explanation, in keeping with other observed asymmetric pattern, features might be biased growth. Model simulations with unbiased growth display limited and dorsoventrally-symmetric net changes in RDVP of melanophores during growth; these data match the original observations poorly (Fig. 8-3B, compare Fig. 8-3B)). However, when we incorporate ventrally biased growth, we observe larger shifts in RDVP, with clear dor-

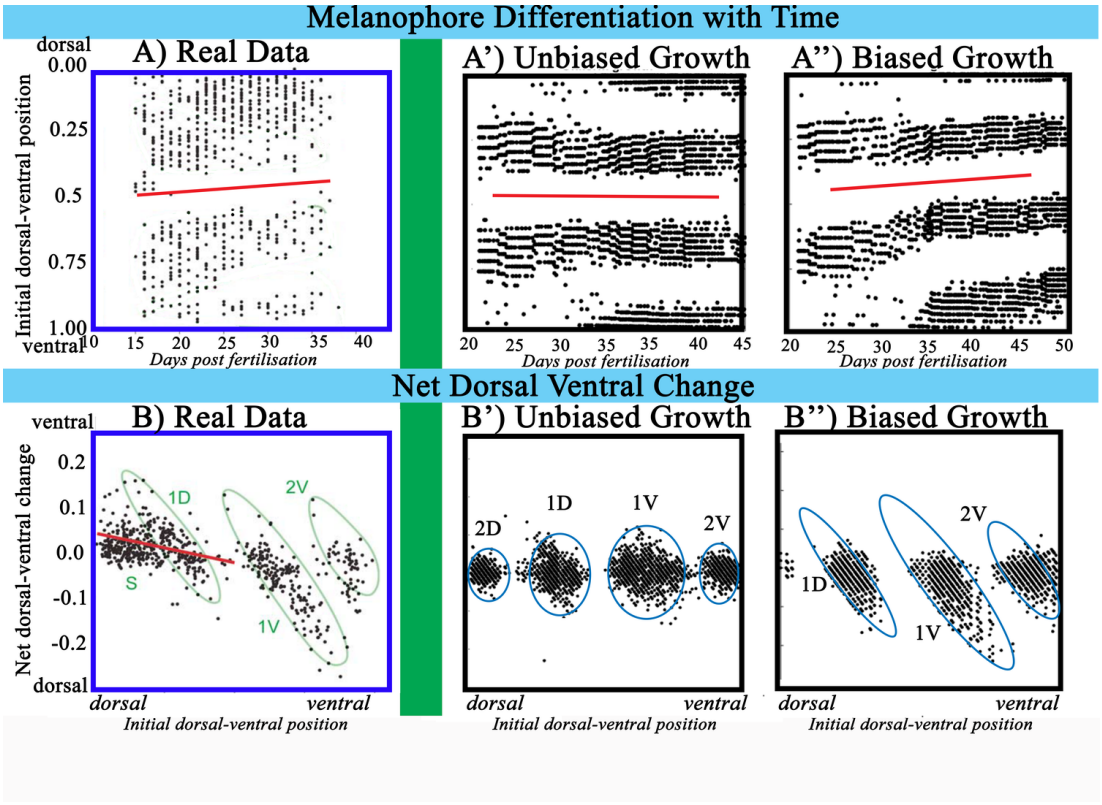


Figure 8-3: *Melanophore births (upper) and movements (lower) in WT fish and simulations of WT fish with and without biased growth. (A) Melanophore differentiation over time as documented by (David M. Parichy and Turner, 2003 [27]). Shown are the initial RDVPs of melanophores first appearing on the day of imaging specified by the x-axis. The red line indicates the predicted centre of X0 interstripe. There is a trend for the positions at which melanophores appear to move dorsally as time progresses. (A') - (A") RDVP of melanophores at time of birth for a representative WT simulations with (A') unbiased growth, (A") biased growth. The relative positions remain symmetric over the DV axis when there is unbiased growth (A'). In contrast a movement trend from ventral to dorsal is observed when there is biased growth (A"). Note that the data in (A) encompasses scale melanophores present at the very dorsal region between stripes 1D and 2D, whereas these are not included in our model of skin pigmentation. For this reason, the differentiation of stripes 1D and 1V melanophore is cleaner in our simulations (A', A") than in the experimental data. (B) Net changes in RDVPs of melanophores (y-axes) are plotted against the initial dorsalventral positions at which these cells were first identified (x-axes) as documented by (David M. Parichy et al., 2003). The red line is drawn for melanophores that appeared between 21 and 26 dpf computed using linear regression and encompasses embryonic melanophores which are not explicitly incorporated in our model. Ellipses are drawn around cells that ultimately contributed to dorsal and scale melanophores (S), the adult primary dorsal melanophore stripe (1D), the adult primary ventral melanophore stripe (1V) and the first adult secondary melanophore stripe at the ventral margin of the flank (2V). Melanophores typically translocate dorsally. The greatest translocation is observed for stripe 1V. (B') - (B") Net changes in dorsoventral positions of melanophores (y-axes) are plotted against the initial dorsoventral positions at which the melanophores first appear (x-axes) over a representative WT simulation with (B') unbiased growth and (B") biased growth. Since net changes are calculated by subtracting the starting position from the final position for each melanophore, net negative changes represent dorsal movements, whereas net positive changes represent ventral movements. Unbiased growth yields no net change on average. In contrast, biased growth yields a net change that is a better representation of the real data. Figures (A) and (B) were reproduced from Parichy et al 2003 [27] and licensed under CC-BY 3.0 (http://creativecommons.org/licenses/by/3.0); published by The Company of Biologists Ltd.*

sal biases, which are most prominent in stripe 1V (Fig. 8-3B), matching closely the key features of the original observations in real fish. We conclude that ventrally biased growth is sufficient to explain the observed relocations of melanophores during metamorphic growth, suggesting that they likely do not result from biased melanophore migration as such.

8.5.4 Biased growth explains asymmetry in missing cell-type and *rose* mutants

Pigment pattern DV asymmetry is also observed in missing cell-type mutant zebrafish, *i.e. pfeffer* (*pfe*), *shady* (*shd*) and *nacre* (*nac*) which lack xanthophores, iridophores and melanophores respectively. These asymmetries are most prominently observed in adulthood and not in juvenile stages [46]. To assess whether these may also be explained by ventrally biased growth, we adapt the bias parameters in our model so that the expected position of X0 at adulthood (RDVP 0.38) is reached by the end of our simulation window at 13.5 mm SL. This acceleration is necessary to overcome the current limit imposed on the duration of our model simulations (terminating in juvenile stages (13.5mm SL)) by a lack information about L-iridophores, which are thought to play a role in pattern maintenance in these later metamorphic stages, although this has not been fully determined [46, 33]. Further discussion and motivation for this adaptation is given in the supplementary material.

We begin by showing that implementing this modified bias does not affect the models recapitulation of the expected WT pattern (Fig. 8-4A-D). Implementing the bias in each of the 'missing cell-type' mutants reveals that biased growth introduces asymmetries to the patterns in a manner qualitatively matching those observed in real fish (Fig. 8-4E-P). In Fig. 8-4H, biased growth within our model results in simulated *pfe* mutant patterns with spot-like features in ventral stripes, contrasting with more connected spots, forming stripe-like features, dorsally. In Fig. 8-4L, biased growth within our model results in simulated shd mutant patterns with ventral melanophore spots being spaced out and rounded, whereas dorsal melanophores spots are more closely spaced, as in real *shd*. In Fig. 8-4P we are able to replicate the expanded X0 interstripe region and bleb-like extensions of dense iridophores and xanthophores seen in real nac (Fig. 8-4P rectangle and Fig. 8-4N). A notable DV pigment pattern asymmetry is seen in rse mutants, in which a reduction in iridophore proliferation generates a partially spotted pattern, with stripes dorsal and ventral to the X0 interstripe, but with more ventral stripes broken into spots. When we reduce iridophore proliferation in simulations to mimic the known biology of rse mutants with unbiased growth we generate a symmetric '*rse*-like pattern [28]: two stripes (often broken into spots) symmetric about the centre of the domain form (Fig. 8-4S). However, it is not until we add biased growth that we can reliably replicate the rse phenotype generating two (largely unbroken) distinct stripes dorsally, with spots more ventrally (Fig. 8-4T).

8.5.5 Biased growth may contribute to the DV asymmetry of Danio nigrofasciatus

Danio nigrofasciatus, a close relative of zebrafish, displays two stripes dorsally and small spots, consisting of a few melanophores more ventrally (Fig. 8-4U,V). We interpret these features as resulting from evolutionary alterations of the zebrafish pigment patterning mechanism as described by [73, 54] (see Methods). Implementing these changes, but without biased growth, our model predicts a DV-symmetric pattern of

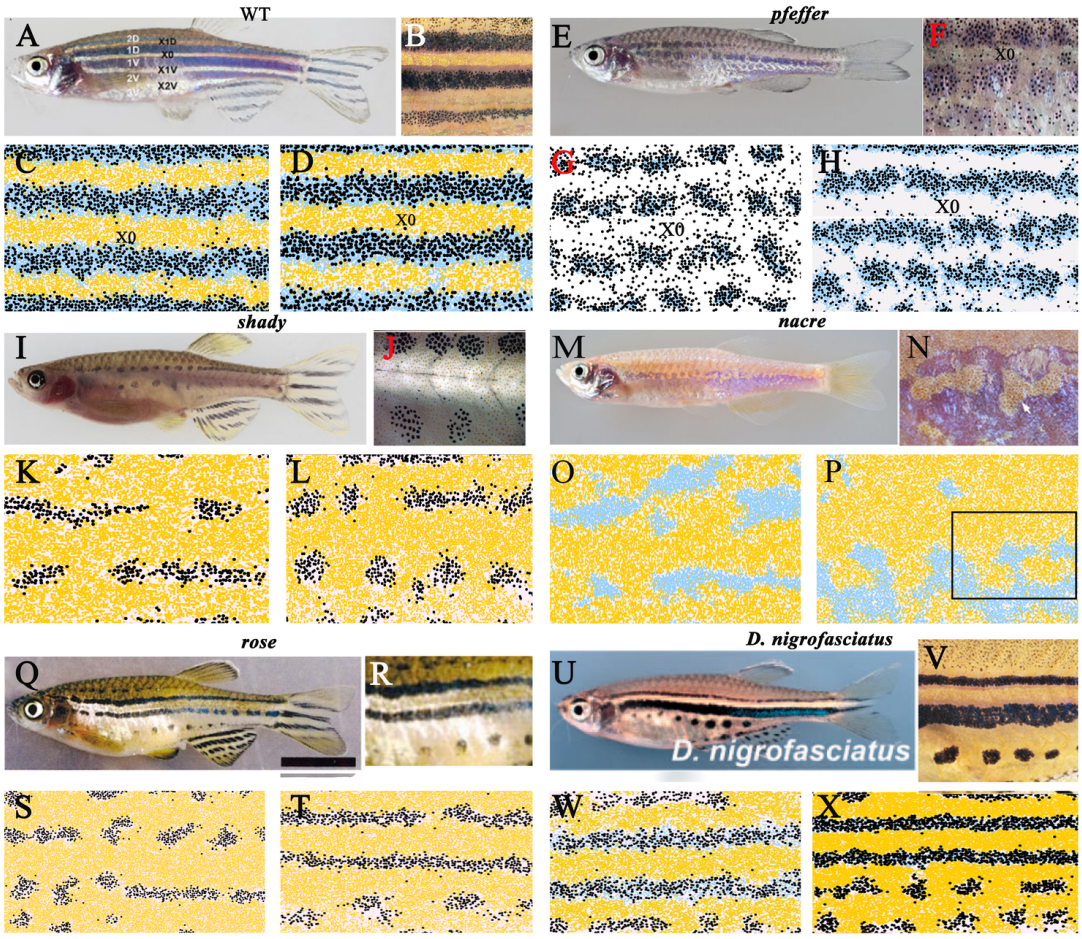


Figure 8-4: *Ventrally-biased growthproduces more realistic patterns for WT, pfe, shd, nac, rse and is able to explain the pattern asymmetry observed in Danio nigrofasciatus. (A), (E), (I), (M). (Q), (U) Images of adult WT, pfe, shd, nac, rse and Danio nigrofasciatus fish respectively. (B), (F), (J), (N). (R) and (V) Close-ups of the adult patterns displayed by WT, pfe, shd, nac, rse and Danio nigrofasciatus respectively. (C), (G), (K), (O), (S), (W) and (D), (H), (L), (P), (T), (X) are the representative simulated patterns of adult WT, pfe, shd, nac, rse and Danio nigrofasciatus at stage J+ with unbiased and biased growth respectively. In each case, simulations performed with biased growth provide better qualitative agreement with the real adult fish than simulations performed with unbiased growth. The box in (P) highlights a region with bleb-like pattern elements repeatedly observed in the simulated nac pattern, and which matches blebs observed in real fish, as shown in (N). Images (A), (B), (E), (F), (I), (J), (M), (N), (U), (V), (Q) and (R) are reproduced from (Frohnhofer et al., 2013 [46]), (Spiewak et al., 2018 [54]) and (Parichy et al., 2000 [55]) respectively and are licensed under CC-BY 3.0 (http://creativecommons.org/licenses/by/3.0); published by The Company of Biologists Ltd.*

four horizontal stripes, with 2D and 2V being broken (Fig. 8-4W). Upon incorporating ventrally biased growth, we reliably (100/100 simulations) recapitulate the asymmetry of the *D. nigrofasciatus* pattern in our simulations, with two stripes dorsally and series of spots more ventrally at simulated stage J+ (Fig. 8-4X).

8.6. Discussion

Here we demonstrate an unexpectedly important role for growth in determining zebrafish body patterns. The importance of growth in influencing development and the final pattern has been suggested by theoretical studies [149, 150, 151, 152, 153, 24, 154, 155] and correlated with inter-specific pattern differences [148], but these have not been tested in the context of real biological data, beyond the effects on specific pattern details such as stripe width [19]. The results of our modelling allow us to reject previous hypotheses that suggested that the iridophore stripe acts as a crucial pre-pattern driving the final pattern orientation. In particular, we show that even drastic changes to the initial iridophore positioning, such as to a vertical stripe, yields relatively small changes to the final pattern. In contrast, dramatic changes in pattern (to maze-like and vertically barred patterns) can be induced by altering the timing and rate of differential growth. We provide evidence that growth may be biased ventrally in zebrafish by showing in silico how it can substantially improve multiple previously unresolved inaccuracies of simulations of both WT and mutant patterns with our model. For example, we show that growth bias in WT Danio rerio provides novel explanations of the dorsal translocation of X0 during development, previously attributed to melanophore differentiation and migration [27], and the ratio of the stripe width of 2D to 2V, as well as improving the overall straightness of X0. X0 is the first stripe to form, so it is likely that maintaining the straightness of this stripe makes the pattern more robust. Growth bias also explains the asymmetries observed in multiple Danio pigment pattern mutants, including the apparently particularly asymmetric rse mutant patterns. *D. nigrofasciatus* also has a striking DV asymmetry to its pattern; by incorporating the proposed changes in patterning mechanism of D. nigrofasciatus (compared to *D. rerio*) as documented by ([73, 54] into our model, and combining this with ventrally-biased growth, we are able to generate the characteristic pattern. Further work is required to determine why we do not generate a 1V stripe thicker than the 1D stripe as seen in *D. nigrofasciatus*; the real pattern may result from persistence of embryonic melanophores, which play an important role in *D. nigrofasciatus* [73], but are not explicitly incorporated in our model. We suggest that growth plays such a significant role in pattern formation for a number of distinct reasons. Firstly, skin growth creates space, allowing for cell differentiation. Indeed this can be seen in the case of the growing angelfish (*Pomacanthus*) which generates new stripes between existing stripes as the developing skin expands [94]. Growth that is biased in a certain region, creates space unevenly, allowing more space for differentiation in some regions, whilst limiting space for differentiation in others; this likely underlies the asymmetries in timing of appearance of stripes 2D and 2V. Secondly, growth can also interrupt cellular interactions by moving cells away from each other. We predict that ventrally-biased growth plays a role in producing the asymmetry of Danio patterns by restricting space for new cells dorsally, thus maintaining the stripes in those regions, whilst creating space ventrally, reducing the effectiveness of cellular interactions and increasing the potential for stripes to break up into spots in those regions. Thirdly, growth can affect the directionality of the pattern by spreading cells along a particular axis, increasing the probability that a pattern becomes biased in that direction. There are some limitations to our approach, principally

due to the current poor understanding of the role of the late appearing L-iridophores in pigment pattern maintenance [33]. For the patterns displayed by mutants, asymmetry in pattern is not qualitatively obvious until a later stage (*i.e.* one where growth has been more extensive) than current experimental understanding permits us to simulate. Thus, in order to replicate the patterns of adult mutant fish we adapted the bias in the model so that we could achieve an X0 position observed at adulthood at an earlier time point than typically observed in real fish. It will be necessary to characterise the patterning role of the L-iridophores and to incorporate this into our model to allow a more complete understanding of the full metamorphic process. A further limitation of the model is that it does not currently include the sites of entry of melanophores onto the flank nor any early differentiating and highly migratory cells. This restricts us from entirely ruling out the concept of migrating melanophores also contributing to the DV asymmetry alongside growth. Nevertheless, our results emphasise the importance of growth and, especially differential growth, of the skin field in determining pigment pattern; growth position, growth rate, and temporal changes in each need to be considered when investigating pattern development of Danio, other fish and vertebrates more generally. Furthermore, such skin growth needs to be considered in reference to the proliferation (cell division) rate of the pigment cells themselves. These mechanisms are likely to be critical in other contexts, including the development of varied pigment patterns within human pathophysiological skin conditions, such as piebaldism [156, 157], where they will act alongside the well-known impact of clonal expansion of pigment cells.

8.7. Methods

8.7.1 Ethics Statement

This study was performed with the approval of the University of Bath ethics committee and in full accordance with the Animals (Scientific Procedures) Act 1986, under Home Office Project Licenses 30/2937 and P87C67227.

8.7.2 Repeated imaging of juvenile fish

We repeatedly imaged N=8 fish every 2 days from 23 dpf for a total of 6 repeats using the confocal microscope. Fish were first anaesthetised in a bath of 0.085% concentration of 2-phenoxyethanol. Once they were non-touch responsive, the fish were removed from the bath and placed in a shallow pool of 0.085% 2-phenoxyethanol and placed under the microscope for imaging. Fish were never anaesthetised for more than 10 minutes. At the end of the experiment, fish were euthanised by Schedule 1 killing.

8.7.3 2-phenoxyethanol

A working solution of 0.085% 2-phenoxyethanol was prepared by diluting 2-phenoxyethanol (Sigma-Aldrich) with system water.

8.7.4 Mathematical model of zebrafish pigment pattern formation

Mathematical model

The model used for this experiment is explained in detail in (Owen et al., 2020) and we summarise its main features here. The model is a comprehensive on-lattice stochastic

mathematical model of zebrafish stripe formation, spanning the developmental period defined by 8 13.5 mm SL. The model consists of three stacked lattices representing the three pigment cell layers as observed in the literature [33]: the melanophore layer accommodating melanophores, an iridophore layer which accommodates loose and dense iridophores and a xanthophore layer which accommodates xanthophores and xanthoblasts. The five pigment cell types are modelled as individual agents which occupy lattice sites on these domains. Volume exclusion rules apply so that no site can be occupied by more than one cell. Cells interact, differentiate, proliferate, die and move at a rate consistent with the biological data (provided this is known). The domain grows over time through the insertion of lattice sites at a rate which mimics the natural growth of fish embryos. Boundary conditions are periodic across the horizontal boundaries and reflecting along the vertical boundaries. We implement periodic boundary conditions along the horizontal axis based on the assumption that the rate at which cells leave along these boundaries is approximately equal to the rate at which cells enter the domain at the opposite side. In practice the choice of reflecting or periodic boundary conditions makes very little difference to the appearance of the pattern. The model is updated according to the Gillespie algorithm [158]. To do this we treat all the potential actions, for example cell birth or domain growth as individual 'events, each with an exponentially distributed waiting time that corresponds to their mean rate of occurrence. To update the model at any given time , an exponentially distributed waiting time is generated until the next possible 'event occurs (based on the rates of all of the possible events). Next a uniformly distributed random number determines which event occurs based on the relative probability of each event. For some events to occur there are also additional constraints that must be met (for example, if a cell is attempting to move, the space into which it is attempting to move must be unoccupied). If conditions required for that event to occur are met, the event is implemented, whereas if they are not, then there is no change. Once an event is accepted, the domain is updated accordingly, and time is also updated by drawing an appropriate exponentially distributed random number. This process repeats until we reach the end of pattern metamorphosis marked by the domain reaching a length of 13.5 mm SL. Previously, we have shown that our model can not only consistently replicate qualitative and quantitative features of WT patterns, but also other zebrafish mutant patterns, demonstrating its validity in investigation of zebrafish pattern formation [28].

Implementing growth bias

In our original model when a growth event in the vertical (without loss of generality) direction is chosen to occur, a site for the growth event is chosen uniformly at random from each column [28]. Supposing the vertical extent of the domain is currently L sites, uniform unbiased growth dictates that the probability that a growth event occurs at a particular site is $\frac{1}{L}$. A new site is inserted below the chosen site with probability 0.5, and above the chosen site with probability 0.5. The corresponding sites in that column are shifted upwards to make room for the new site. We incorporate growth bias by changing the probability that a growth event occurs at a given position, such that it occurs with probability q below γL, where $q \in [0, 1]$ and $\gamma \in [0, 1]$. We determined the values of the bias parameters q and γ for zebrafish by comparing simulated domains against five experimentally measured characteristics described in the supplementary material. Explicitly, we found that bias parameters $q = 0.28$, $\gamma = 0.15$ replicate normal WT growth. In the later sections we employ a stronger bias of $q = 0.5$, $\gamma = 0.25$ to bring about the same shift in X0 at an earlier time point (see supplementary material).

Position of X0

The position of X0 is calculated by first compiling the set of positions and that define the maximum (T) and minimum (B) extent of the X0 interstripe respectively in each column, i. In this way

$$R_i = \frac{T_i - B_i}{2} + B_i, \tag{8.1}$$

the set of positions centred within the X0 interstripe, is a one-dimensional representation of the central interstripe. (The full algorithm for generating R_i is also used to determine the tortuosity of the stripes and is given in detail within the supplementary material of the paper: [28].) The position of X0 in terms of lattice sites is the mean of which we define as . The relative position of X0 is then calculated as:

$$RP = \frac{p - 0.5}{H} \tag{8.2}$$

Where H is the height of the domain.

Stripe straightness

To measure stripe straightness, we first generate a one-dimensional representation (R_i) of the stripe (as described in the above section for the central interstripe). From this line R_i we calculate the stripe straightness SS(x), measured as the ratio of the total length of our line represented by the set of points R_i, which we refer to as R to the straight-line distance between its ends (C).

$$SS(x) = \frac{C}{R} \tag{8.3}$$

For more detail see the appendix of Owen *et al.*, 2020 [28].

Implementation of Danio rerio mutants

shd The gene *shady* (*shd*) encodes zebrafish *leukocyte tyrosine kinase* (Ltk) which plays a role in S-iridophore specification [48, 47, 50]. As a result, strong *shd* mutants lack iridophores. To simulate this defect, we remove all S-iridophores from the initial domain. New S-iridophores are only generated by the proliferation of existing S-iridophores, thus this removes the generation of any S-iridophores throughout the simulation.

nac The gene *nacre* (*nac*) encodes transcription factor Mitfa [51]. *nac* mutants lack melanocytes throughout embryonic and larval development. To simulate this defect, we remove all melanocytes from the initial conditions. We set the rate of new melanophore differentiation to be 0.

pfe Gene *pfeffer* (*pfe*) encodes for *colony stimulating factor 1 receptor (csf1ra)* which plays a role in xanthophore specification and migration [52]. In strong alleles, adult fish exhibit no detectable xanthophores in the body of embryos. To simulate this defect, we remove all xanthophores and xanthoblasts from the initial domain. Since new xanthophores are only generated by the proliferation of existing xanthophores, and the differentiation of xanthoblasts this means that no xanthophores appear on the domain during the simulation.

rse Rose (*rse*), encodes the Endothelin receptor B1a and has been shown to act cell-autonomously in S-iridophores; homozygous mutants result in a reduction of S-iridophores to approximately 20% of that seen in WT (observed in stage PB and adult fish). To simulate *rse* we change the initial conditions so that the number of iridophores on the domain is one fifth of the usual number. We place the remaining iridophores uniformly at random within the space of the central three rows, since some S-iridophores still appear along the horizontal myoseptum in *rse*. We also reduce their rate of proliferation to one fifth of the usual number. For more detail about the simulation of any of these mutants, please see the supplementary material of (Owen *et al.*, 2020, [28]).

8.7.5 Implementation of *Danio nigrofasciatus*

A brief summary of the known developmental differences between *Danio nigrofasciatus* and *Danio rerio* is given by [73, 54]: during pattern metamorphosis, fewer metamorphic melanophores differentiate, and instead a greater number of embryonic melanophores survive. Iridophores do not proliferate as frequently. We interpret this in our model as a decreased rate of differentiation of melanophores in later stages; increased survival rate of melanophores in early stages (as a proxy for the number of embryonic melanophores that survive but are not represented explicitly in our model) and a decreased rate of iridophore proliferation.

8.7.6 Dorsal -ventral relocation of melanophores

To determine the dorso-ventral relocation of melanophores we recorded the position of each melanophore upon appearance and subtracted its final position to determine its net movement.

8.7.7 Pair correlation function

The Pair Correlation Function (PCF) used in Fig. 8-1 is the PCF as described by [105]. We use the PCF on the melanophore domain. The value of the PCF at any given distance x in the horizontal or vertical direction respectively is the number of melanophores on the domain that are exactly this distance apart divided by the number we would expect if the same number of melanophores were places on the domain uniformly at random. If the value of the PCF at a certain distance is 1 this suggests that there is no spatial correlation between cells at that distance. If the value of the PCF is greater than 1 then this suggests there is spatial correlation at this distance. If the value of PCF is less than 1 then this suggests that there is spatial anticorrelation at this distance.

8.8. Statistics

8.8.1 Linear regression

To create the line of best fit in Fig. 8-2B we performed linear regression using the sckit-learn package in python. In particular we used model type LinearRegression from which we determined the intercept and slope of the line.

Supplementary Material

January 28, 2021

1 Context

A juvenile zebrafish grows linearly at an average rate of $33\mu m$ per day in the dorso-ventral (DV) direction. During this time, the position of the X0 interstripe, which is initially in the mid-point of the DV axis moves dorsally by an amount proportional to the height of the fish over time. We suggest that this is caused by a biased growth phenomena whereby, the probability of a growth event occuring is greater ventrally than it is dorsally. In order to investigate this, we build a simple one-dimensional model of the growing DV axis, incorporate growth bias and use the results of this simple model in order to guide the implementation of growth in our comprehensive zebrafish model (described in the main text). This one-dimensional system of ODEs is much less computationally expensive than our full zebrafish model, allowing us to compute the results quickly.

The supplementary material is split into several sections. In Section 2 we derive a general master equation for the probability of occupancy over time for a linearly growing one-dimensional domain. In Section 3 we generate a system of ordinary differential equations for the occupancy of sites on the growing iridophore domain. Next we turn our attention to the position of growth. In Section 4 we incorporate an "unbiased growth" scheme whereby, at each growth event, the position of growth is uniformly distributed across the domain. In Section 5 we build upon the simple unbiased growth linear model to incorporate biased growth. Finally in Section 6 we compute the predicted positions of X0 at time $t = 8$ days using our simple one-dimensional biased linear model with different bias parameters. Comparing the results of these simulations with real data, we are thus able to deduce a subset of parameter values which correspond to real fish biased growth rates and which we test - in the main text - using our comprehensive model.

2 General equation for the probability of occupancy on a growing domain

In this Section we will derive a general equation for a linearly growing one-dimensional domain. Suppose at time t we have a domain of length $\ell(t)\Delta y$ decomposed into sites each of width Δy. Sites are either occupied (with at most a single cell), or unoccupied. Suppose that at time t, sites $\{j_1, \dots j_m\} \in \{1, \dots \ell(t))$ are occupied and all other sites are empty. We can define a density function $\rho_N^L(t)$ as the probability that the length of domain $\ell(t) = L$ and site N is occupied at time t. $\rho_N^L(t)$ is defined for all $L \in \{1, \dots L_{max}\}$ and $N \in \{1, \dots N\}$ where L_{max} is defined upper limit for the length of the domain. At time t:

$$\rho_N^L(t) = \begin{cases} 1 & \text{if } L = \ell(t) \text{ and } N \in \{j_1, \dots j_m\}, \\ 0 & \text{otherwise.} \end{cases}$$

Suppose that we define the rate of growth on this domain to be linear in time with rate λ. When a growth event occurs, a site i on a domain of length L is chosen with probability $b_i^L(t)$. When our discrete domain growth is implemented by inserting sites, we enforce that there is an equal probability that a new site will be inserted before or after the chosen site. In order to investigate how the relative positions of agents on top of the lattice are altered by growth, we can determine an equation for $\rho_N^L(t + \Delta t)$ by considering all movement into and out of this state. For all N and $L < L_{max}$, we can describe the evolution of this system using a simple equation:

$$\begin{aligned} \rho_N^L(t + \Delta t) = &\lambda \left(\sum_{i=N+1}^{L-1} b_i^{L-1}(t)\rho_N^{L-1}(t) + \frac{b_N^{L-1}(t)\rho_N^{L-1}(t)}{2} + \sum_{i=1}^{N-2} b_i^{L-1}(t)\rho_{N-1}^{L-1}(t) + \frac{b_{N-1}^{L-1}(t)\rho_{N-1}^{L-1}(t)}{2} \right) \\ &+ (1-\lambda)\rho_N^L(t) + o(\Delta t), \end{aligned} \tag{1}$$

where $o(\Delta t)$ is the probability that more than one birth event occurs in Δt.

The terms are described as follows. The first term in the equation describes the probability that the domain was length $L-1$, space N was occupied at time t and a growth event occurred after the position N. The second term gives the probability that the domain was length $L-1$, site N was occupied at time t and a growth event occurred at the position N but a site was inserted to the right of site N so its occupancy did not change. The third term gives the probability that the domain was length $L-1$, site $N-1$ was occupied at time t and a growth event occurred before position $N-1$ causing site N to become occupied. The fourth term gives the probability that the domain was length $L-1$, site $N-1$ was occupied at time t and a growth event occurred at position $N-1$ but a new site was inserted at position N, subsequently causing all sites including N to be shifted one place to the right. The final term describes the probability that the domain was length L and site N was occupied at time t and no growth event occurred. We can use this equation to generate a system of ODEs which will determine how $\rho_N^L(t)$ evolves over time.

2.1 Special case: $L = L_{max}$

Despite the fact that in principle a domain could grow arbitrarily large in a finite time, in order to derive a closed system of ODEs we specify a maximum domain size, L_{max}. Once the domain reaches this size, it can never leave this state. Therefore the equation for $p_N^{L_{max}}(t)$ is defined as follows:

$$\rho_N^{L_{max}}(t+\Delta t) = \lambda\Bigg(\sum_{i=N+1}^{L_{max}-1} b_i^{L_{max}-1}(t)\rho_N^{L_{max}-1}(t) + \frac{b_N^{L_{max}-1}(t)\rho_N^{L_{max}-1}(t)}{2} + \sum_{i=1}^{N-2} b_i^{L_{max}-1}(t)\rho_{N-1}^{L_{max}-1}(t) + \frac{b_{N-1}^{L_{max}-1}(t)\rho_{N-1}^{L_{max}-1}(t)}{2}\Bigg) + \rho_N^{L_{max}}(t) + o(\Delta t). \tag{2}$$

3 General set of simultaneous ODEs for the probability of occupancy on a growing domain

We can generate a set of simultaneous ODEs for this model if we rearrange equations (1) and (2) and take the limit as $\Delta t \to 0$. For $L \neq L_{max}$, our general equation (1) becomes:

$$\frac{d\rho_N^L(t)}{dt} = \lambda\left(\sum_{i=N+1}^{L-1} b_i^{L-1}(t)\rho_N^{L-1}(t) + \frac{b_N^{L-1}(t)\rho_N^{L-1}(t)}{2} + \sum_{i=1}^{N-2} b_i^{L-1}(t)\rho_{N-1}^{L-1}(t) + \frac{b_{N-1}^{L-1}(t)\rho_{N-1}^{L-1}(t)}{2} - \rho_N^L(t)\right) \tag{3}$$

3.0.1 Special case: $L = L_{max}$

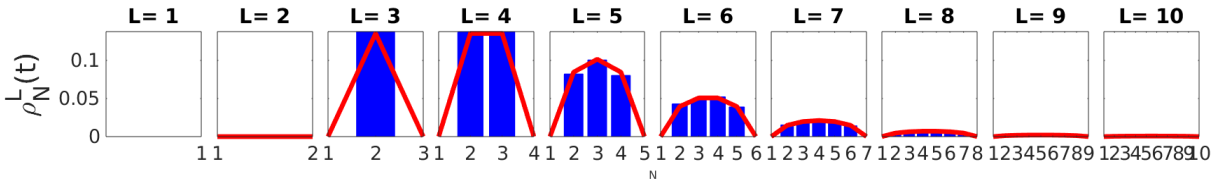
Similarly, if we rearrange equation (2) and take the limit as $\Delta t \to 0$ at the boundary then for $L = L_{max}$ the equation (1) becomes:

$$\frac{d\rho_N^{L_{max}}(t)}{dt} = \lambda\left(\sum_{i=N+1}^{L_{max}-1} b_i^{L_{max}-1}(t)\rho_N^{L_{max}-1}(t) + \frac{b_N^{L_{max}-1}(t)\rho_N^{L_{max}-1}(t)}{2} + \sum_{i=1}^{N-2} b_i^{L_{max}-1}(t)\rho_{N-1}^{L_{max}-1}(t) + \frac{b_{N-1}^{L_{max}-1}(t)\rho_{N-1}^{L_{max}-1}(t)}{2}\right). \tag{4}$$

4 Unbiased growth

Now that we have a general ODE system for describing the occupancy of sites on a growing domain, we can consider different growth bias phenomena by altering term $b_i^L(t)$. To implement unbiased growth, the probability of a site being chosen for a growth event is chosen to be the same across all sites. Therefore:

$$b_i^L(t) = \frac{1}{L} \quad \text{for all } i \in \{1,\ldots L(t)\} \tag{5}$$



Supplementary Figure 1: **Example of using the 'unbiased' growth method.** A visualisation of $\rho_N^L(t)$ at time $t = 5$ where the initial domain size is $\ell(0)$=3 with one site at position 2 occupied, $L_{max} = 10$, and where growth is linear with rate $\lambda = 0.4$. The red line is the solution to ODEs (6)-(7). The blue bars correspond to the occupancies averaged over 5000 runs of the stochastic simulation.

4.1 Unbiased linear growth

To determine a general ODE for unbiased linear growth we substitute equation (5) into equation (1) and simplify to get:

$$\frac{d\rho_N^L(t)}{dt} = \lambda\left(\frac{2L-2N-1}{2(L-1)}\rho_N^{L-1}(t) + \frac{2N-3}{2(L-1)}\rho_{N-1}^{L-1}(t) - \rho_N^L(t)\right). \tag{6}$$

4.2 Special case: $L = L_{max}$

We substitute equation (5) into equation (2) to determine an ODE for our special case when there is unbiased growth.

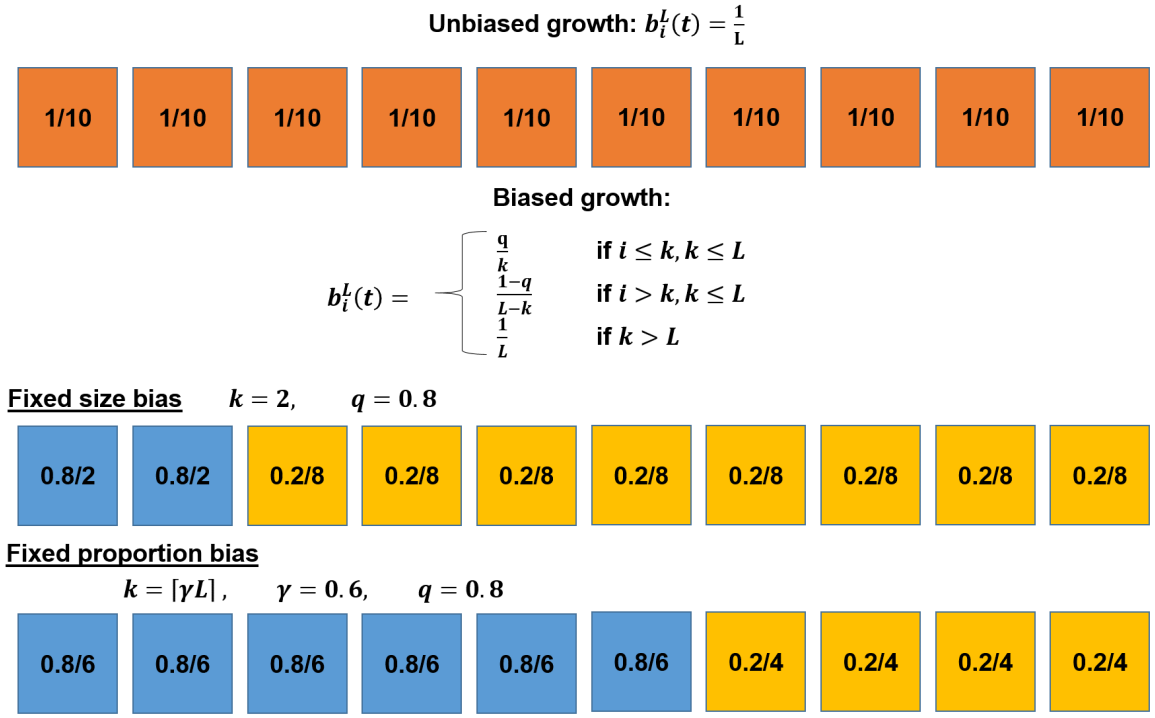
$$\frac{d\rho_N^{L_{max}}(t)}{dt} = \lambda\left(\frac{2L-2N-1}{2(L_{max}-1)}\rho_N^{L_{max}-1}(t) + \frac{2N-3}{2(L_{max}-1)}\rho_{N-1}^{L_{max}-1}(t)\right) \tag{7}$$

4.3 Comparing ODE solutions to stochastic simulations with unbiased growth

We can check that the ODE system is a good representation of the growth mechanism by comparing the solutions of the ODEs with the averaged occupancies generated using stochastic simulation. In Fig. 1, we compare the solutions to ODEs (6)-(7) with the averaged occupancies of 5000 stochastic simulations. We initialise the domain so that that $\ell(0) = 3$ and site 2 is occupied. We evaluate the domain at time $t = 5$, where $L_{max} = 10$ and $\lambda = 0.4$. For the ODE solutions, we start by initialising so that $p_2^3(0)$=1 and $p_N^L(0)$=0 for $N = 1,...L_{max}$ and $L = 1,...L_{max}$. We compute the solutions to ODEs (6)-(7) at time $t = 5$ using the `MATLAB` function `ode15s`. Stochastic simulations are realised as follows. To initialise a simulation we specify a state vector $X \in \{0,1\}^{\ell(0)}$. At time t the state vector represents a domain of length $\ell(t)$ and the occupancy of all of the sites at this time point; $X(N)$=1 if the site N is occupied and $X(N) = 0$ if the site is unoccupied. In this case, at the start of our simulation $X = [0,1,0]$. This represents a domain with initial length $\ell(0) = 3$ with site 2 occupied. We set time $t = 0$. The system is updated as follows.

1. An exponentially distributed waiting time, τ, is generated for a growth event to occur using the Gillespie algorithm.
2. If the time $t = t + \tau$ is greater than $t_{max} = 5$ then the simulation is terminated and we save the final configuration (final size $\ell(5)$ and i such that $X(i) = 1$). Otherwise the simulation continues and a growth event occurs.
3. A position of growth $i \in 1,...L$ is chosen uniformly at random from all the sites.
4. A random number $r_0 \in U(0,1)$ is generated. If $r_0 < 0.5$ a new site is inserted to the left of the chosen site and all subsequent sites, including the chosen site, are moved forwards one position to the right. If $r_0 > 0.5$ then a new site is inserted to the right of the chosen site. All subsequent sites to the chosen site are moved forwards one position to the right.
5. Return to step 1

We run the stochastic simulations 5000 times. Once all 5000 simulations are complete we divide the number of times each double $[i, \ell(5)]$ occurred (where $\ell(5) \in \{1,...L_{max}\}$ and $i \in \{1,...\ell(5)\}$), by the total number of repeats. The averaged occupancies, as well as the solutions to the ODEs, are displayed in Fig. 1. At time $t = 5$, with our



Supplementary Figure 2: **Comparing the probability that a site is chosen as the position of growth using the unbiased and biased implementation for a domain of length 10.** When we implement unbiased growth, the probability that the position of growth is at any domain site is equally spread across all positions (top). Thus the probability of any given site being the position of growth on a domain of length 10 is L =0.1. When we incorporate biased growth but k is a constant (middle) then the probability that one of the sites being chosen for position of growth is at or below k is q/k. For example, if $k = 2$ and $q = 0.8$ then the probability that sites 1 or 2 are chosen as position of growth is $q/k = 0.8/2 = 0.4$. The probability that a site whose index is greater than 2 is $(1-q)/(L-k)$= 0.2/8=0.025. When $k = \gamma L$, k changes with the domain size. Therefore, on a domain of size $L = 10$, sites 1-6 are chosen with probability 0.8/6 and sites 7-10 are chosen with probability 0.2/4.

growth parameters, the domain rarely reaches length 10, therefore the probability of any site N on a domain of length $L = 10$ is close to 0. Furthermore, the domain never shrinks implying that the probability that the domain is length 1 or 2 and time $t = 5$ is exactly 0. The probability that site 2 is still occupied on a domain of length 3, is the probability that no growth event occurred between $t = 0$ and $t = 5$. With unbiased growth and a symmetric initial condition the probability that site N on domain L is occupied is symmetric along the axis. This means that it is always most likely that the initially central site, remains central, regardless of the domain length L. As expeceted there is no bias using this method.

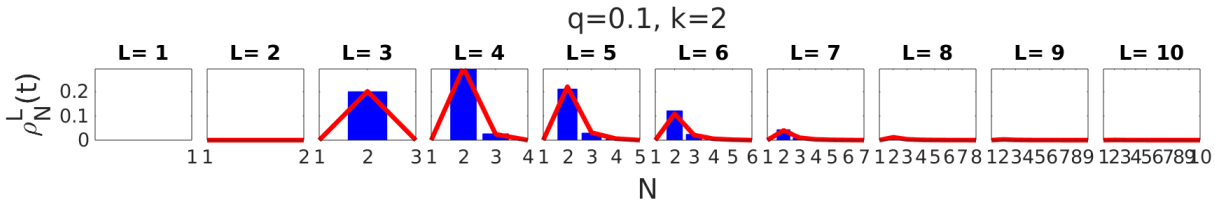
5 Biased growth

We can also choose other values for $b_i^L(t)$ so that growth is biased towards one side of the domain. We consider two forms of biased growth:

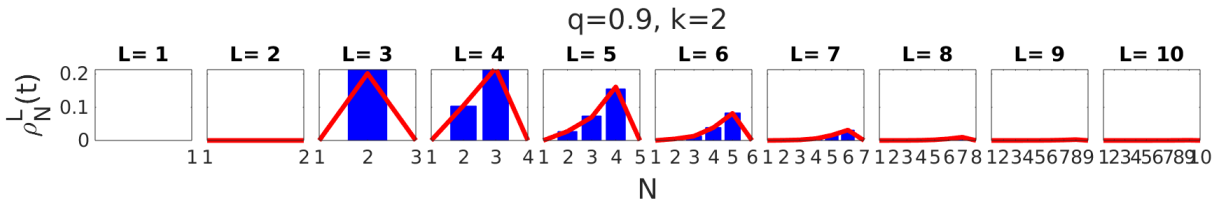
1. **Fixed-size bias:** The domain grows with probability $q \le 1$ at sites whose index is less than or equal to $k \in \mathbb{Z}$, $k \le L(0)$ and with probability $1-q$ at sites whose index is greater than k.

2. **Fixed-proportion bias:** The domain grows with probability $q \le 1$ at sites whose index is less than or equal to $k(t) = \lceil \gamma L(t) \rceil$ where $\gamma < 1$ and with probability $1-q$ at sites whose index is greater than k.

In these cases the growth position bias is given by:



(a)



(b)

Supplementary Figure 3: **Example of using the fixed-size bias growth method for different values of q where k is a constant.** A visualisation of $\rho_N^L(t)$ at time $t = 5$, where the initial domain size is $\ell(0)$=3 with one site at position 2 occupied, $L_{max} = 10$, $k = 2$ and growth events occur at a constant rate $\lambda = 0.4$. (a) and (b) show the results using different q (a) $q = 0.1$, (b) $q = 0.9$. Each subplot in panels (a) and (b) from left to right displays the probability that the domain is length $L = 1, ...L_{max}$ and the site N on the domain is occupied at time t. The red line represents the solution to ODEs (8)-(9) a time $t = 5$. The blue bars correspond to the occupancies averaged over 5000 repeats of the stochastic simulation.

$$b_i^L(t) = \begin{cases} \frac{q}{k} & \text{if } i \leq k, k \leq L \\ \frac{1-q}{L-k} & \text{if } i > k, k \leq L \\ \frac{1}{L} & \text{if } k > L. \end{cases}$$

For fixed-size bias, k is a constant, whereas for fixed-proportion bias, $k = \lceil \gamma L \rceil$, which changes over time. A comparison of the implementation of biased and unbiased growth is given in Fig. 2. Fixed-size bias could correspond to a growth scenario in which a specific group of cells in a certain location divide with one frequency while other cells divide with another. Fixed-proportion bias corresponds to a situation in which growth is biased towards a fixed proportion of the domain. Substituting the birth rate equation (5) into equation (6) and evaluating we get the following when $N > 1$ and $L > 1$ gives:

$$\frac{d\rho_N^{L(t)}}{dt} = \lambda \left(\frac{(2L-2N-1)(1-q)}{2(L-k-1)} \rho_N^{L-1}(t) + \left(q + \frac{(1-q)(2N-3-2k)}{2(L-k-1)} \right) \rho_{N-1}^{L-1}(t) - \rho_N^L(t) \right). \tag{8}$$

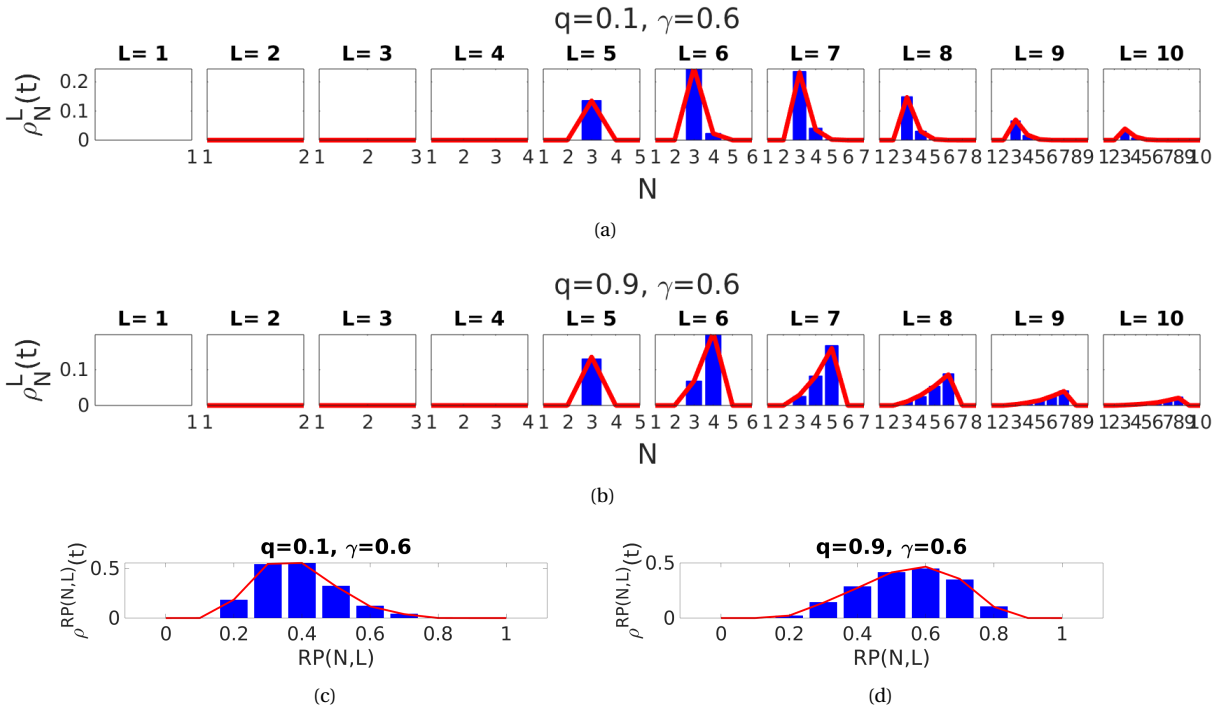
5.1 Special case : $L = L_{max}$

If we rearrange equation (2) and take the limit as $\Delta t \to 0$ at the boundary then for $L = L_{max}$ the equation (1) becomes:

$$\frac{d\rho_N^{L_{\max}}(t)}{dt} = \lambda \left(\frac{(2L_{\max}-2N-1)(1-q)}{2(L_{\max}-k-1)} \rho_N^{L_{\max}-1}(t) + \left(q + \frac{(1-q)(2N-3-2k)}{2(L_{\max}-k-1)} \right) \rho_{N-1}^{L_{\max}-1}(t) \right). \tag{9}$$

5.2 Comparing ODE solutions to stochastic simulations for fixed size biased growth

Let us consider the case where k is a constant. In Fig. 3 we compare two values of q for when $k = 2$ and the initial condition is a size $\ell(0) = 3$ domain with site 2 occupied at time $t = 0$ (and all other sites are unoccupied). When q is small, the probability of a growth event is much more likely for the sites beyond the currently occupied site. Therefore the probability that sites greater than 2 become occupied is small. When q is large, the probability of a growth event moving the currently occupied site to the right is high hence the probability that sites greater than 2 become occupied by time $t = 5$ is high.



Supplementary Figure 4: **Example of using the fixed-proportion bias growth method for implementing growth for different values of q.** A visualisation of $\rho_N^L(t)$ at time $t = 5$ where the initial domain size is $\ell(0)$=3 with only the site at position 2 occupied, $L_{max} = 10$, $\gamma = 0.6$ and where growth events are chosen to occur a constant rate $\lambda = 0.4$. Panels (a) and (b) show the results using different q, (a) $q = 0.1$, (b) $q = 0.9$. Each subplot in panels (a) and (b) from left to right displays the probability that the domain is length $L = 1, ...L_{max}$ and the site N on the domain is occupied at time t. The red line represents the solution to ODEs (8)-(9) at time $t = 5$. Panels (c) and (d) display the same results as panels (a) and (b) where the position of occupancy has been normalised with respect to the length of the domain using equation (10) and occupancies average. The red line is the solution to ODEs (8)-(9). The blue bars correspond to the occupancies averaged over 5000 runs of the stochastic simulation.

5.3 Comparing ODE solutions to stochastic simulations for fixed proportion biased growth

Now let us consider fixed-proportion bias: $k = \lceil \gamma L \rceil$. In Fig. 4 we compare the probability of occupancy using two values of q for when $\gamma = 0.6$ and the initial domain size is $\ell(0) = 3$ with only site 2 occupied.

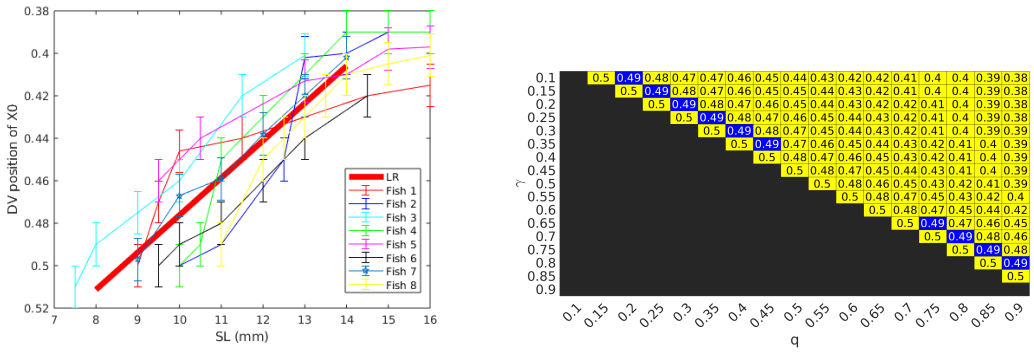
6 Predicting zebrafish growth

6.1 Comparing real and model-generated data

When we take measurements in real fish, we look at the position of X0 relative to the total domain size. In order to be able to compare our theoretical results with real data, we compute the relative position with respect to domain size. The relative position of site N on a domain of length L is given by:

$$\mathrm{RP}(N, L) = \frac{N - 0.5}{L}. \tag{10}$$

The subtraction of 0.5 from N, computes the position of the centre of the lattice site N. We divide by the total height of the domain to find the position relative to the domain size. In Fig. 4 (c)-(d) we display the same results as in Fig. 4(a)-(b) except the position of occupancy is normalised so that it is relative to the domain size (RP). When $\gamma = 0.6$ and $q = 0.1$ an initially central site is most likely to end up around 0.3 by time $t = 5$ (Fig. 4(c)). In contrast, when $\gamma = 0.6$ and $q = 0.9$, an initially central site is most likely to end up around 0.6 by time $t = 5$ (Fig. 4(d)).



(a) (b)

Supplementary Figure 5: **The position of the X0 interstripe moves dorsally as the fish grows bigger and we can incorporate bias in our model to incorporate this change.** Panel (a) shows how the position of X0 along DV axis varies with SL (mm). The thick red line represents the linear regression of the points. The equation of this line is given in equation 6.3. The other coloured lines represent the position of X0 along the DV axis against SL (mm) for all of the sampled fish 1-8. Relative DV X0 positions were calculated post imaging by dividing the position of the centre of the X0 interstripe by the full length of the fish. Fish were imaged on their sides, however, the exact angle of which they faced onto the glass varied. For this reason, determining the very top and very bottom of the fish was difficult and could not be measured exactly. For this reason, we estimate that we were able to determine the relative position to be correct to the nearest 1% thus error bars indicate the precision error of each point. (b) Using the one-dimensional model we can determine suitable q and γ to test in the comprehensive model. For example, using the red line in (a) we can determine that by time $t = 12$ days the position of X0 should be at approximately 0.49 on the relative DV axis (see Section 6.4). The heatmap shows the predicted position of X0 at time $t = 12$ using different values of q and γ in the one-dimensional model. Values highlighted in blue are those that match the real data (0.49) best.

In our model of zebrafish growth we will use fixed-proportion growth bias. We choose this as we estimate it to be the most biologically realistic. In this section we use our one-dimensional model to determine parameters q and γ for use in our comprehensive zebrafish model. Our aim is to find suitable q and γ such that the position of X0 is a close fit to the real data.

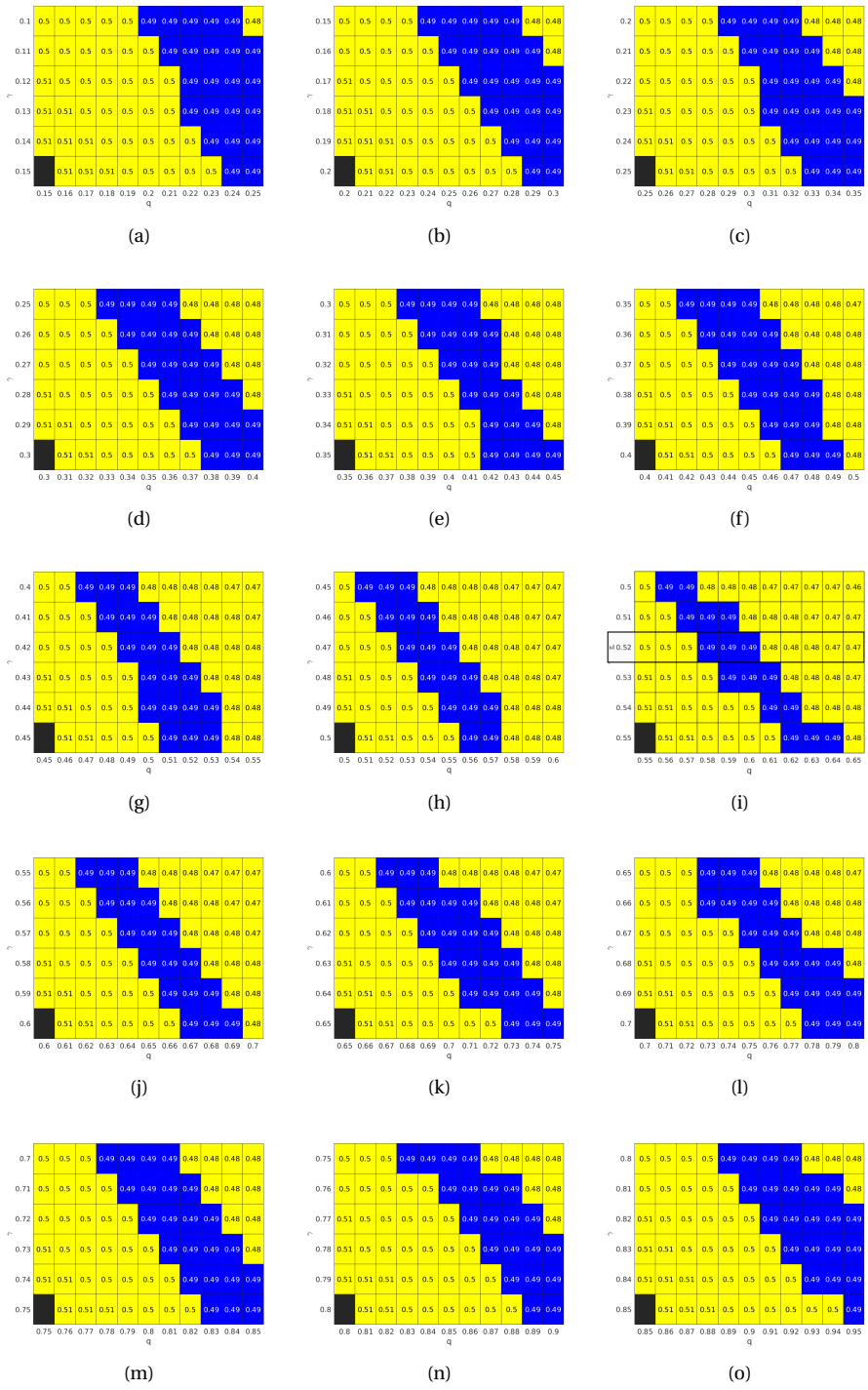
6.2 Initialising our one-dimensional model

In the comprehensive 2-dimensional model, we model the iridiphore layer at time t on the body of the fish as a two dimensional lattice model of height $\ell(t)$ and width $W(t)$. Each square lattice site of size $20\mu m$ is either occupied with a single iridophore (loose or dense type) or is unoccupied. At the start of the simulation, the domain (which corresponds to a fish that is 8 mm Standard Length (SL) in length), has a height of $\ell(0)$=50 lattice sites, and a length of $W(0)$=100 lattice sites. This corresponds to the full height of the fish and 2mm of the length. The initial iridophore interstripe is incorporated as three rows of dense iridophores in the middle of the fish such that the central row is at DV position 0.5. In our one-dimensional model, we represent only the vertical (DV) axis of the fish. Therefore $\ell(0) = 49$ and we represent the centre of the domain as a single occupied site at position 25.

6.2.1 Computing the DV position

In our comprehensive model, row 1 represents the most ventral part of the body and row represents the most dorsal part. Similarly position 1 in our one-dimensional model represents the most ventral region and $\ell(t)$ represents the most dorsal region. In order to compute the relative DV position of a site (i, j) chosen at random on our comprehensive model, or equivalently the DV position of site i on our one-dimensional model we make the following calculation:

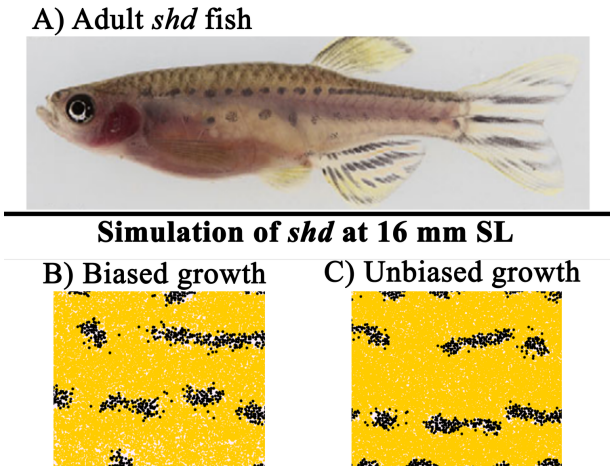
$$DV(i) = 1 - \mathrm{RP}(i)., \tag{11}$$



Supplementary Figure 6: **A fine-grained search of parameters reveals more parameter possibilities.** An extensive search of the parameter space local to the q and γ that yield positions of X0 highlighted in blue in the heatmap given in Fig. 5(b). The heatmap shows the predicted position of X0 at time $t = 12$ days for each q and γ. Values coloured in blue indicate matches to the real data (i.e. the position of X0 is predicted to be 0.49)

6.3 Initialising the position of X0

In Fig. 5(a) we compare the position of X0 along the DV axis with SL in mm. Using a linear regression model we compute that the equation of the best-fit line is



Supplementary Figure 7: **Using the bias parameters on fish shd at a later time point of 16 mm SL.** (A) The adult *shd* mutant. (B) Representative simulation at 16 mm SL using the bias $q = 0.28, \gamma = 0.15$. Spot-like features are more likely ventrally than dorsally. (C) Representative simulation of *shd* at 16 mm SL using unbiased growth. Melanophore spots are elongated dorsally and ventrally.

$$f(x) = mx + c \tag{12}$$

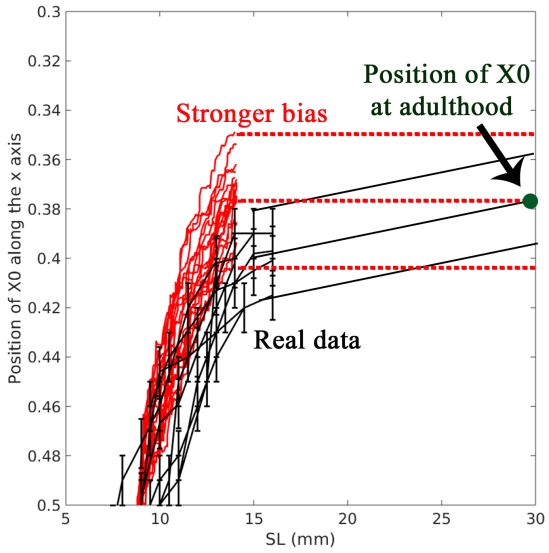
where $m = -0.018$ and $c = 0.652$ to 3 decimal places (dp) and where $f(x)$ represents the relative position of X0 and x represents SL. The line generated from this equation is given by the red line in Fig. 5(a). Using this equation we can compute that at 8 mm SL the relative position of X0 is 0.5 to 2 dp. This matches the position of X0 initialised in our comprehensive and 1-dimensional model.

6.4 Calibrating our data

We wish to compare the position of X0 in our one-dimensional model at time $t = 12$ days with the data we have about real fish. To do this, we must first calibrate the speed of growth in the real data with the speed of growth that is implemented in our comprehensive model. Our comprehensive model is calibrated such that the rate of growth is 0.13 mm SL per day on average. To calibrate the data we first determine the predicted length of the fish at time $t = 12$ using the growth rate implemented in our comprehensive model. If a fish is 8 mm SL at time $t = 0$ and the fish is growing at a rate of 0.13 mm per day, the fish should be 9.56 mm SL at time $t = 12$. Using equation (6.3) the relative position of X0 on the DV axis 12 days later of a fish growing at a rate of 0.13 mm per day is 0.49 to 2dp.

6.5 Using the one-dimensional model to compute q and γ

We initialise our one-dimensional model as described in Section 6.2. We set the growth rate λ (*i.e.* the rate of addition of boxes) to be $\frac{33}{20} = 1.65$ per day. This is because the rate of growth in the vertical direction in zebrafish is 33μm per day and each box is 20μm wide. Therefore the rate at which a new row is added to the domain (of size 20) is 33/20 = 1.65 per day. Using the ODEs (8)-(9) we solve until time $t = 12$ and compute the predicted position of X0 at this time for different values of q and γ, taking the position of modal probability at time $t = 12$ and using equation (11) to compute the DV position. In Fig. 5(b) we display a heatmap of the predicted position of X0 at time $t = 12$ rounded to 2dp using our one-dimensional model as described. The values range between 0.38 and 0.5. Since our real data indicates that at time $t = 12$ the relative position of X0 is approximately 0.49 we highlight those values in dark blue. However, this mesh is coarse and there may be other suitable values of q and γ that may give a similarly good match in unexplored regions of the parameter space. We expand our parameter sweep using the one-dimensional model for parameter values within 0.05 of the coarse parameter pairs ($[q, \gamma]$) that generate a DV position of 0.49. The results are given in Fig. 6. In doing so, we determine 156 values of q and γ that generate an X0 position that is 0.49 to 2 dp (highlighted in blue).



Supplementary Figure 8: **Determining a suitable bias for predicting adult patterns.** Red lines represent 20 simulations with bias parameters $q = 0.5$ and $\gamma = 0.25$. Black lines are the real data points. Adult fish of length 30 mm SL exhibit an X0 position around 0.38. Extrapolating from the experimental data, we expect that X0's relative DV position moves dorsally, linearly between 15 mm SL and 30 mm SL. Lack of experimental evidence around the role of L-iridophores precludes us from simulating from 13.5 mm SL. The bias extracted from the real data allows us to get simulations that are correctly biased up to 13.5 mm SL. Using the stronger bias $q = 0.5$ and $\gamma = 0.25$ we can achieve the same final X0 position but within the time frame for which the assumptions of our underlying model are valid.

We generate 20 repeats for each of the 156 distinct (q, γ) pairs. For each repeat we generate time-series data about the position of X0 (y_i) and SL (mm) (x_i) at time point i. We score each $[q, \gamma]$ pair based on a number of factors: the error in the position of X0 when compared with the real time series data given by equation

$$\sum_{i=1}^{N} \frac{|f(x_i) - y_i|}{N}, \tag{13}$$

where the function f is as defined in equation (6.3). We also collect the following data about the 8 fish: the ratio of 1D to 1V at the final time point (which should be 1); the ratio of 2D to 2V at final time point (which should be close to 0.5); the final position of 1D (close to 0.2) and; the final position of 1V (which should be close to 0.6). Thus for each repeat we compare our expected outcome with the real data. For each data point, if the error meets a tolerance within the range of 10 percent of the real value then its score for this category is 1, otherwise it is zero.

The combination of $q = 0.28$ and $\gamma = 0.15$ scores the highest across all categories. We implement $q = 0.28$ and $\gamma = 0.15$ in Fig. 2 of the main paper. The position of X0 with SL is given in Fig. 2G of the main text and is a good match to the real data between 8 mm SL and 13.5 mm SL.

7 Predicting patterns of adult mutant fish

The simulated pattern at 13.5 mm SL with the bias defined by $q = 0.28$, $\gamma = 0.15$, displays X0 at position at approximately 0.44 relative DVP as in real WT fish at 13.5 mm SL (Fig 5(a)). We can simulate other mutant *Danio rerio* up to 13.5 mm SL using this bias. Mutant fish with very obvious asymmetry include missing cell mutants: *nac*, *pfe* and *shd*. Unfortunately, for these mutants the asymmetry becomes more obvious only later in development. For example, whilst juvenile *nac* mutants display some asymmetry in the pattern by 13.5 mm SL (such as the dorsal shift of X0), the more obvious asymmetries (including the contrast of homogeneous xanthophores and iridophores dorsally compared with blebs of xanthophores and iridophores in the ventral

region (Fig. 2 of the main text)) are seen at adulthood. We predict that this is due to the continuation of biased ventral growth from 13.5 mm SL to 28 mm SL. Thus is evidenced by looking at the relative DV position of X0 in adult fish which is closer to 0.38.

There is currently a lack of literature about the role of cell type: L-iridophore. Between 13.5 mm SL and 16 mm SL, L-iridophores appear, which are not included in the model and which likely play a role in stripe maintainance. For this reason, our model is currently not valid for zebrafish that exhibit L-iridophores between 13.5 mm - 16 mm SL. The only mutation that is an exception is the *shd* fish. The mutant fish *shd* does not produce iridophores and thus for *shd* we can simulate to 16 mm SL. In Fig. 7 we show that by biasing growth such that $q = 0.28$ and $\gamma = 0.15$ until the fish is 16 mm SL generates a pattern that is asymmetric. The spots appear more elongated dorsally and rounded ventrally. This demonstrates that our prediction is robust. To investigate missing cell mutants *nac* and *pfe*, however, we must take a different approach. Since we can only simulate between 8 mm SL and 13.5 mm SL we choose (in a similar method as before) a bias that can yield an average position of X0 of 0.38 by 13.5 mm SL. In this way, we are able to achieve the same final X0 relative DV position, as at adulthood within the time frame that we are able to simulate in. The X0 relative DV position against SL with $q = 0.5$ and $\gamma = 0.25$ is given in Fig. 8. We use this bias to simulate WT and mutant zebrafish in Fig. 3 and 4 of the main text.

8 Summary

To summarise, we generated a 1-dimensional model to investigate bias as a simpler and quicker alternative to our comprehensive model of zebrafish pattern formation. This allowed us to find bias parameters which matched the real growth of fish from 8 mm SL to 13.5 mm SL. We demonstrated that this worked well for replicating the adult pattern of the *shd* mutant which lacks L-iridophores (as well as for replicating the pattern of J+ WT fish - see main text). However, the bias we determined did not work well for replicating the adult patterns of WT fish and mutants due to the lack of experimental data about L-iridophores. To combat this, we generated bias parameters which could give the same RDV position of X0 at adulthood within the time in which we can simulate. Finally, we used these bias parameters in the main text and use these to show that we are able to replicate the asymmetries in patterns suggesting that they are caused by biased growth ventrally.

8.9. Conclusions

In this chapter we explored the impact of growth on zebrafish stripe formation. The most important findings of this paper are the dramatic changes to the pattern that can be elicited by simply changing the growth rates as well as the finding that growth biased to occur more frequently in the ventral region during pattern formation is the most parsemonious explanation for the dorso-ventral asymmetric patterns seen in *Danio rerio*. These findings will be important both for considering pattern development in other animals with more extreme growth rates, as well as when considering the genetics of mutants of *Danio rerio* in the future.

In the context of the thesis, this chapter demonstrates the capability of our mathematical model to investigate hypotheses which are either too difficult or impossible pursue experimentally, such as the impacts of changing the position or rate of growth. This study also provides a good example of the way in which our mathematical model, alongside biological experimentation, can be implemented in order to uncover details about zebrafish pattern development. The applicability of our model to quantitative analysis is a strength which we hope will be utilised in many future studies.

Chapter 9

Final conclusions and outlook

The main object of this thesis has been to understand the generation of patterns in the animal kingdom using zebrafish as a model organism and mathematical modelling as a tool. As such, this thesis presents a series of studies in which, by employing an individual-based modelling technique, we investigate previously unexplained phenomena regarding the pattern forming process in zebrafish.

In Chapter 3 we set the context of the thesis by providing a general introduction to the pattern forming process in zebrafish, paying particular attention to all known individual cellular interactions. We further provided a summary of zebrafish pigment pattern mutants with hypotheses relating gene functionality with phenotype. In Chapter 4 we reviewed the literature of mathematical models of zebrafish pattern formation, splitting them into component-level models – models which only consider components of the overarching system and comprehensive models – models which attempt to represent the full system. We showed that of the two, comprehensive models had more scope for generating testable predictions and testing biological hypotheses relating to zebrafish pattern formation. The findings of these reviews motivated the use of a comprehensive mathematical modelling approach in order to understand zebrafish stripe formation. In anticipation of building our own comprehensive on-lattice agent based mathematical model of zebrafish pattern formation, in Chapter 5, we generated a statistical tool which we could use to inform and validate our model by comparing it directly to experimental data. In particular, we generated a new PCF for discrete domains that could identify and quantify spatial patterns. We showed that our PCF improved on previous PCFs for discrete domains as well as generalising the PCF to domain discretisations in higher dimensions. We built our comprehensive mathematical model of zebrafish pattern formation in Chapter 6. Incorporating all the known and relevant biology of zebrafish pattern formation determined from the literature, we were able to develop an individual based model that could replicate pattern formation both in WT and mutant zebrafish. Crucially we demonstrated that the known cellular interactions along with some we had to hypothesise are indeed consistent with zebrafish pattern generation. We subsequently used our model to answer long-standing biological questions. For example, we determined a minimal set of rules that could reproduce the *leo* mutant pattern as well as using the model to perform a brief investigation into the functionality of growth in the pattern forming process.

In the second part of the thesis we shifted our focus towards using the model alongwith experimental results to explore the previously unexplained phenomena wherein the adult *Danio* patterns are dorso-ventrally asymmetric, despite the assumed symmetry of all known cellular interactions. Inspired by the prospect of being able to generate data for our model regarding, for example, the degree of dorso-ventral pattern asymmetry, in Chapter 7 we provided a thorough investigation into juvenile zebrafish anaesthesia. Our aim was to provide a suitable protocol for repeatedly anaesthetising and imaging fish during the juvenile stage encompassing pattern metamorphosis. Our findings showed that juvenile zebrafish were more sensitive to anaesthetic than fish in other stages and that 2-phenoxyethanol at concentration 0.085% is the most suitable anaesthetic for repeatedly imaging fish at this juvenile stage, provided that we maintain

the fish on a cottonwool pad doused with the anaesthetic. In Chapter 8 we returned to our investigation into dorso-ventral asymmetry. Based on the prediction that asymmetry is caused by growth being more rapid ventrally than dorsally, we adapted the model to incorporate growth bias parameterising the bias using results of repeatedly imaged fish (making use of the method given in Chapter 7). Using the parameterised model we showed that biased growth can explain the asymmetry of pattern formation in WT and mutant fish as well as other previously unexplained phenomena such as the migration of melanocytes dorsally. Our results further demonstrated that growth plays a greater role in determining pattern type than the initial iridophore stripe contradicting previous theories.

Overall, the aim of this thesis was to build a mathematical model of zebrafish pattern formation that can be used in conjunction with experimentation to aid future understanding of the pattern forming process. As expected of any investigation, the research presented in this thesis left several questions unanswered. In the following section we will highlight some avenues for further investigation.

9.1. Future Research

9.1.1 Derivations of continuum models of zebrafish pattern formation

An advantage of using an on-lattice model for studying zebrafish pattern formation is its ammenability to the derivation of a continuum model. A benefit of developing a continuum model is that we can perform analytical studies. Whilst developing a continuum model for the full zebrafish pattern model would be infeasible due to the sheer complexity of the system, it would be possible to generate continuum model for say the interactions of iridophores and melanophores on *pfe* for example which has reduced number of interactions. To generate such a model we would construct a conservation of occupancy statement for a site on the melanophore layer, $\mathbf{M}(i, j)$, by considering all transitions into and out of that site between time t and $t + \delta t$ as in Simpson *et al* [159]. Similarly we could construct a conservation of occupancy statement for site on the iridophore layer: $\mathbf{I}(i, j)$. Upon rearranging the two equations and taking limits [159], we could derive a continuum model describing the averaged occupancy over time for the melanophore and iridophore layer.

Future work would be to generate a continuum model for *pfe* based on our comprehensive model which we could then use to test whether or not pattern formation in the case of *pfe* could be described as a Turing process.

9.1.2 Gene analysis

In Chapter 6 we used the model to determine the minimal rules for *leo* mutant based on observations of *leo*, *leo;nac*, *leo;shd* and *leo;pfe*. We showed that upon implementing these hypotheses in our model, not only could we reproduce all aforementioned mutant patterns but we could also replicate the patterns produced by mutant *leo;sbr*. These results suggest that our hypotheses are very promising, however, without real data indicating so, we cannot be certain our hypotheses are correct. Indeed, since then, other in depth investigations into the *leo* mutation have been made, including the generation of mutants that lack functionality of *leo* in melanophores and xanthophores only [71]. This was considered in the discussion of Chapter 6 where we showed that our hypothe-

ses *could* explain the pattern changes observed for these mutants, suggesting that our hypotheses were valid. In a similar way, future work will be required using our model in conjunction with other experimental studies to determine whether or not our hypotheses about *leo* are correct. Our *in silico* findings from our hypotheses also provide testable predictions for the future. For example in the discussion of Chapter 6 we predicted that *Tg(mitfa-cx39.4)cx39.4*$^{-/-}$;*shd* would display a reduction in melanophores with no obvious spots unlike *shd*. Related to this, as more data becomes available about other cell-communication mutants such as *luc* and *obe*, we will be able to hypothesise genetic function using our model of these too.

9.1.3 Extending the model beyond juvenilehood

In Chapter 8 we saw that one major limitation of our model is that it is only valid for predicting pattern development for fish aged between 21 dpf and 50 dpf, or more accurately, between developmental stages PB and J+. This proves a problem for studying any pattern features that appear later than this period. For example, in Chapter 8 we saw that asymmetry in the patterns of some mutants such as *shd* and *nac* were only prominent in adult development and not at juvenile stages. Since lack of biological information on the behaviour of L-iridophores precluded us from modelling adult development, we were forced to alter the bias in the growth in order to see the same changes in the required simulation time.

In order to extend the model, future work will be required to investigate the role of L-iridophores. L-iridophores first appear on the body of the fish around stage J+ below the melanocyte layer and likely play a key role in pattern maintenance between stage J+ and adulthood [33]. Once more data is available about how L-iridophores interact with other cells, we will be able to incorporate them explicitly, allowing us to simulate even further forwards in time and to explore other aspects of the pattern forming process.

9.1.4 Dorso-ventral countershading

The adult stripe pattern is not the only pigment pattern to be observed on the adult zebrafish. Superimposed onto the striping pattern, is another evolutionarily derived dorso-ventral countershading (DVC) pattern [9, 8] described as a dark dorsum and light ventrum. Unlike striping, this pattern is a highly conserved evolutionary trait ubiquitous to most animals and fish that are subject to light. For a full and detailed review of this pigment patterning mechanism as well as all known genetics related to it see Chapter 3, Section 3.17.

In a preliminary study, as part of a collaboration with Dr Josep Rottlant from the Aquatic Molecular Pathobiology Lab in Vigo, Spain, we began to investigate dorso-ventral countershading. Preliminary findings are given in Appendix A. In particular we accumulated data pertaining to the distribution of pigment cells on scales of mutants without ligand *asip1*: (*asip1* KO), Transgenic *asip1* fish (*asip1* Tg) as well as melanocortin receptor mutants MC1R KO, and MC2R KO. Our major findings were that: dorso-ventral countershading is related to dorsal and ventral scale cell composition (Fig A-1 and consistent with other studies [8]); that melanocyte and xanthophore distributions dorso-ventrally across the mutants vary and often non-intuitively (Fig A-3) and; that iridophores appear in different shapes dorsally and ventrally (Fig A-4).

Future work would be required to incorporate these findings related to dorso-ventral countershading into our working model of zebrafish body pattern formation. Dorso-ventral countershading does have subtle effects on the body pattern of the zebrafish,

however, it is largely an effect of differential scale pigmentation. Therefore, we would first focus on developing a model describing the differentiation of melanocytes on the scales of the zebrafish. The melanocortin receptors are G protein-coupled receptors [160]. Thus in a method similar to that of Bridge *et al* [161] who also model the interactions of G protein-coupled receptors, based on predictions made in Section A.3, we would generate a series of ordinary differential equations that describe the concentrations of melanocortin receptors, *asip1* and other ligands over time. Due to the fact that activated melanocortin receptors effect the differentiation of melanophores, we could then relate our model to our real data about pigment cell generation on the scales. In an iterative process, we would work closely with biologists, generating experimentally testable predictions (such as predictions about the phenotype of other cross-mutants) to guide future research and using the results of experiments to guide our modelling. Once satisfied with our model for scales, we would test our predictions by attempting to recreate the effects of DVC on the stripes of the zebrafish using our model in Chapter 6: for example,*asip1* KO fish display an extra interstripe ventrally [8]. To do this we would incorporate our rules about the interactions of *asip1* and the melanocortin receptors to body pattern formation. Our results would build a much more complete understanding of zebrafish pattern formation.

9.1.5 Beyond zebrafish

Though the comprehensive model we built in Chapter 6 is aimed to develop our understanding of zebrafish specifically, in the future we aim to use the model to understand other related processes outside of zebrafish. We have already touched upon this albeit briefly in this thesis. We showed in Chapter 8 that we could use the model to explore the evolution of *Danio* by investigating the relationship between *Danio rerio* and *Danio nigrofasciatus*. In Chapter 8 we saw that our model can be applied more generally to explore aspects of growth with relation to a complex patterns in many fish. More generally still, the techniques used in this thesis might also form the basis of many other investigations into animal patterns.

In summary, this thesis represents a mathematical modelling approach to understanding zebrafish stripe formation. We generated a comprehensive mathematical model that can be used to analyse features of zebrafish pattern formation. From a general perspective, the studies presented in this thesis demonstrate the power of a mathematical modelling approach to allow a deeper understanding of complex biological phenomena.

Appendices

Appendix A: Dorso-ventral countershading: A preliminary study

Superimposed onto the zebrafish striping pattern, is an evolutionarily derived dorso-ventral countershading (DVC) pattern [9, 8] described as a dark dorsum and light ventrum. Unlike striping, this pattern is a highly conserved evolutionary trait [162] ubiquitous to most animals that are subject to light. Whilst there are largely differing explanations for its function and its reason for being conserved [3] in fish and some mammals, the pattern is thought to be a basic camouflage mechanism from predators or prey positioned directly above or below the animal. From above, fish exhibiting DVC, camouflage with the dark sea and from below, the bright sun.

DVC is achieved differently between mammals and fish. Mammals have only one type of pigment producing cell type, a melanocyte, which can generate two pigment types; yellow pheomelanin or dark brown eumelanin. The DVC mechanism in mammals is regulated by two loci, the agouti locus and the extension locus which regulates the pigment production. The extension locus encodes for the α-Melanocyte simulating hormone (α-MSH), which binds to the Melanocortin Receptor 1 (Mc1r) on melanocyte plasma membrane inducing cAMP production [163, 164]. The agouti locus encodes for the Agouti Signalling Peptide protein (ASIP) which antagonises this interaction. In particular, binding of agouti lowers the ratio of dark eumelanin to light pheomelanin produced in the melanocytes [165], inhibits melanoblast differentiation and proliferation [166, 167] and promotes dedifferentiation of melanocytes *in vitro* [168]. It is the sustained expression of ASIP in the ventral region which is responsible for a switch in melanocytes from dark eumelanin production to light coloured pheomelanin, whilst the dorsum is more darkly coloured due to the eumelanin [169, 170].

The DVC patterning mechanism differs significantly in fish. In particular, unlike mammals, countershading in fish is achieved by a patterned distribution of the pigment cells, with the light-absorbing and light-reflecting chromatophores mostly distributed in the dorsal and ventral areas, respectively [11, 79]. Despite this, some mechanisms are conserved. For example, dorso-ventral pigmentation gradients in fish is thought to depend on a dorso-ventral expression gradient of *asip1* (the fish ortholog of mammalian ASIP) [8, 88] for the following reasons.

The protein *asip1* is associated with lightly pigmented regions, for example, *asip1* is associated with the light ventrum of non-teleost spotted gar and teleosts; goldfish (*Carassius auratus*), sole (*Solea senegalensis*) and turbot (*Scophtalmus maximus*) and also in zebrafish where *asip1* has been shown to be expressed in a dorso-ventral gradient. Overexpression of *asip1* causes lightening of the pigment pattern, in the case of turbot and sole, injection of *asip1* causes lightening of pigment in this region. In zebrafish, *asip1* overexpression inhibits melanogenesis in the dorsal region ([8, 10] by antagonizing melanocortin signalling (binds to receptor, preventing the effect of other ligands, without causing any activity) [171], but also by promoting the proliferation of dorsal iridophores [8]. Finally, *asip1* is typically associated with the stages of development in which pigment patterns form. In zebrafish, *asip1* expression is largely associated with larvae-adult pattern metamorphosis at 15-30dpf. These results indicate that *asip1* still plays a role in DVC, by inhibiting melanogenasis and by differential chromatophore differentiation.

This analysis leaves an important question concerning the signaling mechanism that *asip1* uses in zebrafish to regulate pigmentation. To date, six different melanocortin receptors (Mc1r-Mc5r) have been characterized in zebrafish. These are Mc1r, Mc2r, Mc3r, Mc4r, Mc5ra and Mc5rb. An agonist is a chemical that binds to a receptor and activates the receptor and an antagonist is a chemical that binds to a receptor and thus blocking or dampens a biological response by blocking it. An inverse agonist is a ligand that binds to the same receptor as an agonist but induces a response opposite to that of the agonist. All of the aforementioned melanocortin receptors bind with Melanocyte-stimulating hormone (MSH) (a likely agonist with varying affinities) and appear to be antagonised by *asip1* [88]. *asip1* has been shown to work as a competitive antagonist of both Mc1r and Mc4r [88, 172] but also as an inverse agonist of the constitutively activated Mc1r ([172, 164]), with low Mc1r activity occurring in the absence of MSH or the presence of *asip1*. Other agonists include Adrenocorticotropic hormone (Acth) which activates Mc1r, and the endogenous antagonists of the agouti signaling protein family (Asip1, Asip2, Agouti-related peptide 1 (Agrp1) and Agrp2).

In this preliminary study we focus on DVC caused by pigment pattern formation on the scales. In particular we will quantify pigment cell contributions on dorsal and ventral scales of the following fish types: *asip1* KO which lacks *asip1* [87], Tg(Xla.Eef1a1:Cau.Asip1)iim05 [8] abbreviated to *asip1* Tg which overproduces *asip1*; Mc1r KO fish which do not produce Mc1r [9] and; Mc2r KO fish which do not produce Mc2r [9]. We compare the number of pigment cells on the scales with WT fish (TU strain). Additionally we compare the pigment producing cells observed on scales of cross mutant fish of the above such as Mc1r KO x *asip1*.

Previous studies of *asip1* have largely quantified *asip1* activity based on the numbers of pigment cells in the dorsal and ventral region without distinguishing specifically those cells in the scales versus those in the skin [8]. Meanwhile a few studies have commented on the vast changes in scale pigmentation between mutant and transgenic fish for *asip1* KO, Mc1r compared to that in wild type [8, 9]. Scales develop around the time of pigment pattern metamorphosis starts; around 9.1-10 SL [82, 83] and their pigment composition appears to have a large impact on the overall dorso-ventral countershading pattern.

Here we present a study of scale pigmentation across wild-type, *asip1* and two melanocortin receptors including Mc1r mutant fish. We demonstrate disruptions of the dorso-ventral pigmentation polarity in these mutant fish, as well as interesting observations regarding Mc1r KO which allude to *asip1* as an inverse agonist for some potential melanocortin receptor repressed by Mc1r. The work presented here was performed in collaboration with Dr. Josep Rottlant at the Aquatic Molecular Pathobiology Lab in Vigo, Spain.

A.1. Materials and Methods

A.1.1 Fish

Fish of the following genotypes were used: TU strain (Tubingen, Nusslein Volhard Lab), Mc1r KO [9], Tg (Xla.Eef1a1:Cau.Asip1)iim05 [8], *asip1*KO [87]. Double transgenic/ mutant lines were obtained by setting up crosses between the respective mutants or transgenic fish.

A.1.2 Imaging

For microscope imaging, zebrafish of 180 dpf were anaesthetised with tricaine methasulfonate (MS-222, Sigma-Aldrich) Scales were isolated using forceps and immersed in drops of a 500ml prepared saline solution consisting of 13ml of 5M NaCl, 1.35ml of 1M KCl, 2.8ml of D-Glucose 1M, 1ml of 0.5M EDTA, 2.5ml of 1M Tris HCl and 479.55ml of milli water. Scales and fish were photographed with a Leica M165FC stereomicroscope equipped with a Leica DFC310FX camera.

A.2. Results

A.2.1 Dorso-ventral countershading is related to cell composition in the dorsal and ventral scales

In Figs A-1(A), (C) and (B), (D) we demonstrate the same WT fish body and head respectively, before and after the scales are removed. We observe that when the scales are removed from the body the fish shows a paler appearance, reflecting the importance of scale pigment cells in determining the DVC appearance (Fig A-1(A)-(B)). This is most prominant in the dorsal region where we observe a strong reduction of melanocytes in the very dorsal region not associated with the striping when the scales are removed (Fig A-1(B) and (D)). This finding is consistent with previous studies that have also suggested that dorso-ventral countershading are related to dorsal scale and ventral scale cell composition in the scales of zebrafish [8]. For example, in wild-type fish, dorsal scales consist of all three pigment-producing cell types, in contrast, ventral scales lack all pigment-producing cell types, consistent with DVC patterning [8].

A.2.2 Melanocyte and xanthophore counts vary dorso-ventrally across all fish types investigated

In Fig A-2 we define the positions from which scales must be plucked in order to be considered dorsal (D), associated with stripes 1D, 2V or to be considered ventral (V). In Fig A-3 we study the relationship between the number of melanocytes and xanthophores respectively and their dorso-ventral position. We found that in all fish types, the number of melanocytes were significantly exaggerated in the dorsal region, with typically over 100 melanocytes per scale more in the dorsal region than anywhere else (Fig A-3, left panel). Moving from 1D to 2V to V (*i.e.* ventrally not including position D) we found that in all but two fish types (WT, *asip1* KO, Asip1-Tg, Asip1-Tg x *asip1* KO, Mc1r KO x *asip1* KO, Mc2r KO) melanocyte numbers were negatively correlated with position ventrally, consistent with the DVC patterning. In the two remaining fish types: Mc1r KO and Mc1r KO x Asip1-Tg, the opposite was true. Moving ventrally from 1D to 2V to V, the melanocyte number increased.

In the case of xanthophore counts in the scales, only four fish types exhibited negative correlation with position ventrally (moving from D to 1D to 2V to V, Fig A-3, right panel). These fish types were WT, Asip1-Tg x *asip1* KO, Mc1r KO x *asip1* KO and Mc2r KO. In all other fish types, (*asip1* KO, Asip1-Tg, Mc1r KO, Mc1r KO x Asip1-Tg and *asip1* KO) the number of xanthophores followed a U-shaped distribution whereby moving ventrally (from D to 1D to 2V to V) numbers decreased to 0 at either 1D or 2V and then increased again moving ventrally to V. In *asip1* KO, Asip1-Tg, Mc1r KO and *asip1* KO the xanthophore number was zero at position 2V. In Mc1r KO x Asip1-Tg the xanthophore number reached zero at 1D.

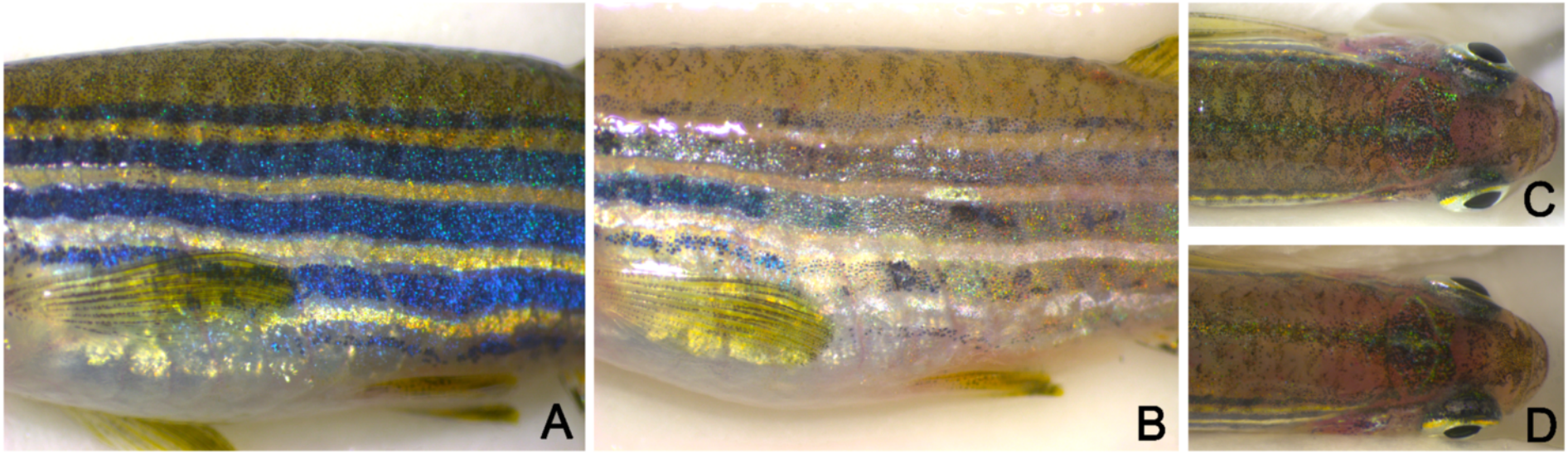


Figure A-1: ***Dorso-ventral countershading is related to cell composition in the dorsal and ventral scales.*** *(A)-(D) Images of the same WT fish (A),(C) with scales and (B),(D) without scales. (A) and (B) display a view of the body, (C) and (D) displays the very dorsal part of the fish.*

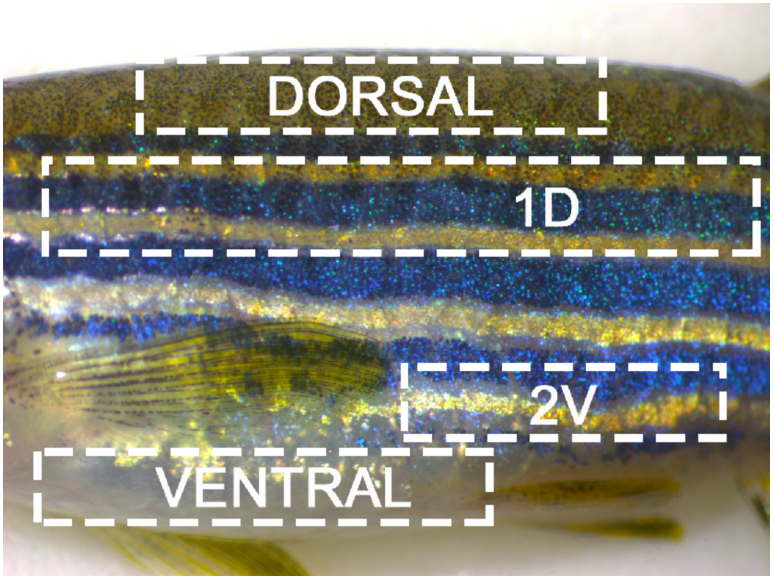


Figure A-2: ***An illustration of the positions from which the scales were extracted.*** *The dorsal and ventral regions are abbreviated to 'D' and 'V' in the text.*

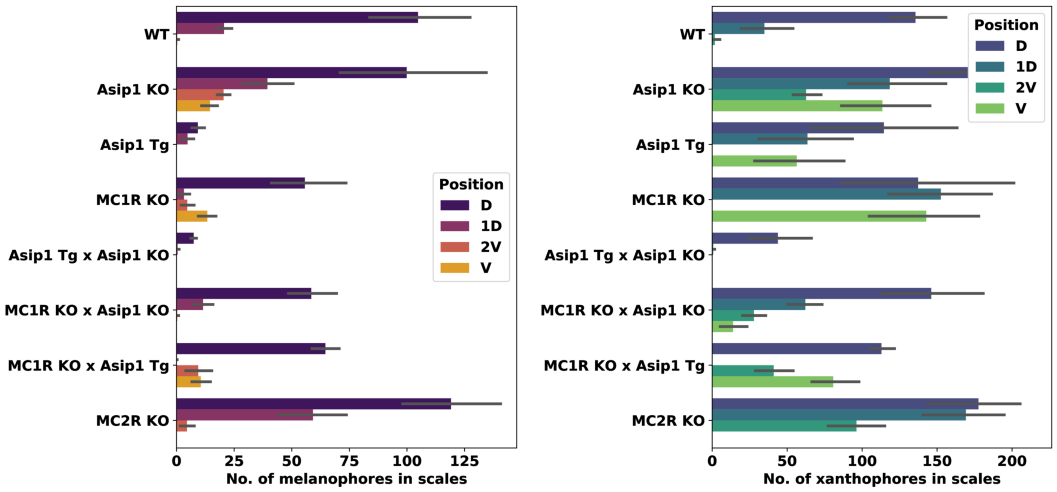
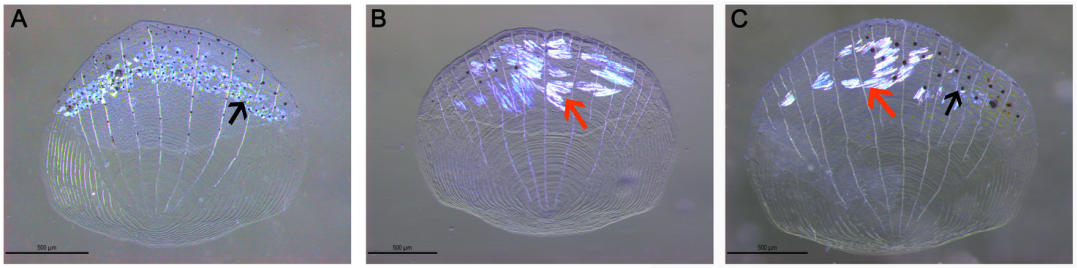


Figure A-3: ***Melanocyte and xanthophore counts vary dorso-ventrally across all fish types investigated.*** *The number of (left) melanocytes, (right) xanthophores in the scales by fish type and position*

A.2.3 Iridophores on the scales appear in two different shapes

We observed that iridophores appeared in two different configurations on the scales, a condensed-type shape which exhibited a small surface area and which were highly associated with melanocytes (Fig A-4(A)). The second was a stratified-type whereby iridophores were rod-like in appearance and were usually found in low melanocyte density regions (Fig A-4(B)). We found that in most cases, any given scale would exhibit just one of the two types, however, some scales exhibited both types simultaneously. In these cases, condensed type iridophores were associated with melanocytes on one half of the scale and the other half were filled with stratified type iridophores (Fig A-4(C)).

Different iridophore shapes were more common in different positions dorso-ventrally (Fig A-4(D)). For example, in the most dorsal region, iridophores were exclusively condensed type across all mutant and transgenic fish. Moving ventrally, in scales from



D Table of iridophore shape type of scales in different positions on the body.

Genetics of fish / Scale position	Dorsal	1D	2V	Ventral
WT	Condensed	Mix	Stratified	None
asip1 KO	Condensed	Mix	Condensed	Condensed
asip1 Tg	Condensed	Mix	Stratified	None
asip1 Tg x *asip1* KO	Condensed	Stratified	Stratified	None
MC1R KO	Condensed	Stratified	Stratified	None
MC1R KO x *asip1* KO	Condensed	Stratified	Stratified	None
MC1R KO x *asip1* Tg	Condensed	Stratified	Stratified	None
MC2R KO	Condensed	Mix	Mix	None

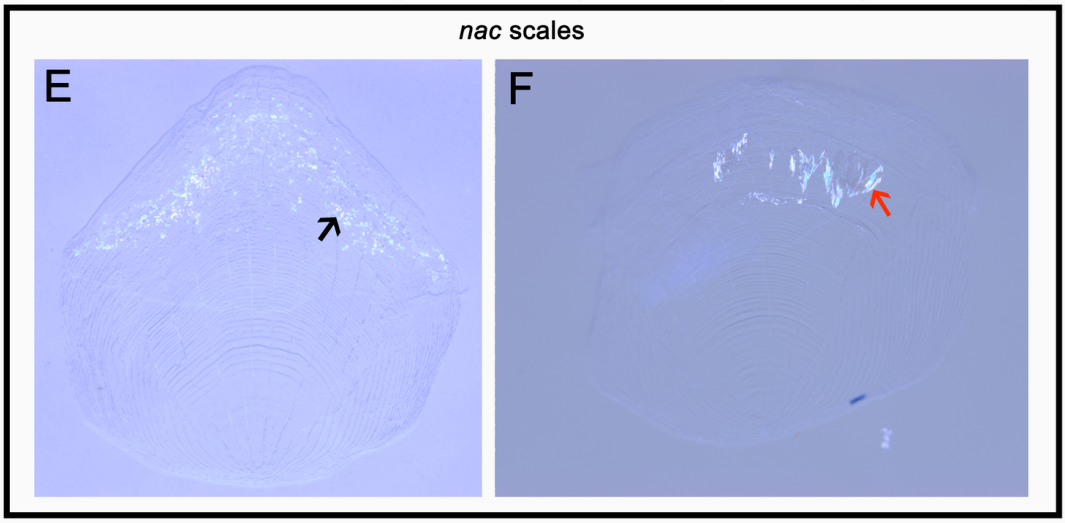


Figure A-4: ***Iridophores on the scales*** *(A)-(C) Scales extracted from WT fish (A) Example of a scale with only condensed-type iridophores. (B) Example of a scale with only stratified-type iridophores. (C) Example of a scale with both condensed- and stratified-type iridophores. (D) Table indicating the type of scales associated with different regions of body for fish of different types. (E)-(F) Scales extracted from nac fish which lack melanocytes. (E) Example of a nac fish scale extracted from the dorsal region. Iridophores form in condensed-type and there are no xanthophores. (F) Iridophores form in a stratified-type and there are no xanthophores. Black arrows indicate condensed-type iridophores. Red arrows indicate stratified-type iridophores.*

1D and 2V, all fish types except *asip1* KO exhibited iridophores in a stratified shape, with some fish types exhibiting an intermediate 'mixed' stage in which some scales exhibited both condensed and stratified iridophore types. At the most ventral part, all fish types except for the *asip1* KO fish lacked iridophores. In *asip1* KO fish, scales consisted of condensed shaped iridophores in all positions except for position 1D where there were also some stratified iridophores.

Finally, since iridophores in the dorsal region that took a stratified shape were typically associated with melanocytes, we studied scales in *nac* mutant fish to determine whether melanocytes caused the shape change. The mutant *nac* lacks melanocytes

throughout embryonic and larval development and thus would also lack melanocytes on the scales [51]. We found that in *nac* fish iridophores still appeared in two types, suggesting that iridophore shape change is not associated with melanocyte proximity (Fig A-4(E)-(F)). We also found that despite *nac* being able to generate xanthophores, there were no xanthophores on any of the scales extracted.

A.3. Predictions

A.3.1 An increase in *asip1* is associated with a decrease in melanocyte production (when all receptors are present)

Our results suggest firstly that *asip1* reduces the production of melanocytes, presumably by antagonising melanocortin receptors Mc1r, Mc4r and/or others for the following reasons. In WT fish, *asip1* is expressed in dorso-ventral gradient, with low expression dorsally and high expression ventrally. In the scales of WT fish we observe no melanocytes ventrally. This is in contrast to *asip1* KO which observes some melanocytes ventrally. In Asip1-Tg mutants, the number of melanocytes is reduced significantly across the domain in comparison to WT fish.

A.3.2 *asip1* is not the only ligand responsible for the dorso-ventral gradient of melanocyte production

In WT fish which can produce *asip1*, there is a steep gradient of high melanocyte production dorsally to no melanocytes ventrally. This is presumed to be caused by the production of *asip1* ventrally which inhibits melanocortin receptors and thus the production of melanocytes ventrally. If *asip1* is the sole ligand involved in DVC we would expect that *asip1* KO fish do not exhibit a gradient of melanocyte production. In contrast, whilst a dorso-ventral gradient in the case of the *asip1* KO is more gradual, this gradient is maintained. This suggests that another ligand alongside *asip1* is responsible for maintaining this DVC.

A.3.3 In the absence of Mc1r, *asip1* is a potential agonist for another melanocortin receptor

Consider the number of melanocytes as we move dorso-ventrally through positions 1D to V for mutants: Mc1r KO, Mc1r KO x *asip1* KO and Mc1r KO x Asip1-Tg in Fig A-3. In Mc1r KO, there is an increase in melanocyte number correlated with an increase in *asip1* transcription. This is also observed in the case of Mc1r x Asip1-Tg. It is not, however, observed when Mc1r KO is crossed with *asip1* KO suggesting that *asip1* is responsible for agonising another unknown receptor responsible for melanocyte production.

A.4. Conclusion

In this short preliminary study we investigated DVC in relation to pigment cell production on the scales. We first showed that without scales the effects of DVC are lessened, indicating that scale pigmentation is important to the DVC pattern. This is consistent with previous studies [9]. We next compared the pigment producing cells present on the scales of WT with fish types that lacked *asip1*, overproduced *asip1* (Asip1-Tg) as well as knockout mutants of Mc1r and Mc2r. We showed that the distribution of pigment

producing cell types varied across all the fish types investigated in an often unintuitive way. Based on these findings we made a set of predictions about the mechanisms behind DVC. Future mathematical investigation combined with focussed experimental work is required to determine whether or not these predictions reflect reality.

Bibliography

[1] C. Li, A. Ji, and Z. Cao, "Stressed Fibonacci spiral patterns of definite chirality," *Appl. Phys. Lett.*, vol. 90, no. 16, 2007.

[2] P. D. Shipman and A. C. Newell, "Phyllotactic patterns on plants," *Phys. Rev. Lett.*, vol. 92, no. 16, 2004.

[3] G. D. Ruxton, M. P. Speed, and D. J. Kelly, "What, if anything, is the adaptive function of countershading?," *Anim. Behav.*, vol. 68, no. 3, pp. 445–451, 2004.

[4] J. L. Kelley, G. M. Rodgers, and L. J. Morrell, "Conflict between background matching and social signalling in a colour-changing freshwater fish," *R. Soc. Open Sci.*, vol. 3, no. 6, 2016.

[5] P. Haffter, J. Odenthal, M. C. Mullins, S. Lin, M. J. Farrell, E. Vogelsang, F. Haas, M. Brand, F. J. Van Eeden, M. Furutani-Seiki, M. Granato, M. Hammerschmidt, C. P. Heisenberg, Y. J. Jiang, D. A. Kane, R. N. Kelsh, N. Hopkins, and C. Nüsslein-Volhard, "Mutations affecting pigmentation and shape of the adult zebrafish," *Roux's Arch. Dev. Biol.*, vol. 206, no. 4, pp. 260–276, 1996.

[6] U. Irion, H. G. Frohnhöfer, J. Krauss, T. Çolak Champollion, H. M. Maischein, S. Geiger-Rudolph, C. Weiler, and C. Nüsslein-Volhard, "Gap junctions composed of connexins 41.8 and 39.4 are essential for colour pattern formation in zebrafish," *eLife*, vol. 3, p. e05125, 2014.

[7] A. Fadeev, J. Krauss, H. G. Frohnhöfer, U. Irion, and C. Nüsslein-Volhard, "Tight junction protein 1a regulates pigment cell organisation during zebrafish colour patterning," *eLife*, vol. 2015, no. 4, pp. 1–25, 2015.

[8] R. M. Ceinos, R. Guillot, R. N. Kelsh, J. M. Cerdá-Reverter, and J. Rotllant, "Pigment patterns in adult fish result from superimposition of two largely independent pigmentation mechanisms," *Pigment Cell and Melanoma Res.*, vol. 28, no. 2, pp. 196–209, 2015.

[9] L. Cal, P. Suarez-Bregua, J. M. Cerdá-Reverter, I. Braasch, and J. Rotllant, "Fish pigmentation and the melanocortin system," *Comp. Biochem. Physiol. -Part A : Mol. Integr. Physiol.*, vol. 211, no. January, pp. 26–33, 2017.

[10] R. Guillot, R. M. Ceinos, R. Cal, J. Rotllant, and J. M. Cerdá-Reverter, "Transient Ectopic Overexpression of Agouti-Signalling Protein 1 (Asip1) Induces Pigment Anomalies in Flatfish," *PLoS ONE*, vol. 7, no. 12, pp. 1–10, 2012.

[11] J. T. Bagnara and J. Matsumoto, "Comparative Anatomy and Physiology of Pigment Cells in Nonmammalian Tissues," *Pigment. Syst. Physiol. Pathophysiol. Second Ed.*, vol. ISBN 0-19, no. -509861-7, pp. 11–59, 2007.

[12] H. G. Frohnhöfer, S. Geiger-Rudolph, M. Pattky, M. Meixner, C. Huhn, H. M. Maischein, R. Geisler, I. Gehring, F. Maderspacher, C. Nüsslein-Volhard, and U. Irion, "Spermidine, but not spermine, is essential for pigment pattern formation in zebrafish," *Biol. Open*, vol. 5, no. 6, pp. 736–744, 2016.

[13] T. E. Woolley, P. K. Maini, and E. A. Gaffney, "Is pigment cell pattern formation in zebrafish a game of cops and robbers?," *Pigment Cell and Melanoma Res.*, vol. 27, no. 5, pp. 686–687, 2014.

[14] S. Dini, "Spatial Quantification and Mathematical Modelling of Tissue Development," 2017.

[15] T. E. Woolley, "Pattern production through a chiral chasing mechanism," *Phys. Rev. E*, vol. 96, no. 3, pp. 1–16, 2017.

[16] S. Kondo, "An updated kernel-based Turing model for studying the mechanisms of biological pattern formation," *J. Theor. Biol.*, vol. 414, no. July 2016, pp. 120–127, 2017.

[17] J. M. Bloomfield, K. J. Painter, and J. A. Sherratt, "How Does Cellular Contact Affect Differentiation Mediated Pattern Formation?," *Bull. Math. Biol.*, vol. 73, no. 7, pp. 1529–1558, 2011.

[18] A. Nakamasu, G. Takahashi, A. Kanbe, and S. Kondo, "Interactions between zebrafish pigment cells," *Pnas*, vol. 106, no. 21, pp. 8429–8434, 2009.

[19] A. Volkening and B. Sandstede, "Modelling stripe formation in zebrafish: An agent-based approach," *J. R. Soc. Interface*, vol. 12, no. 112, p. 20150812, 2015.

[20] D. Bullara and Y. De Decker, "Pigment cell movement is not required for generation of Turing patterns in zebrafish skin," *Nat Commun*, vol. 6, no. May, p. 6971, 2015.

[21] R. Asai, E. Taguchi, Y. Kume, M. Saito, and S. Kondo, "Zebrafish Leopard gene as a component of the putative reaction-diffusion system," *Mech. Dev.*, vol. 89, no. 1-2, pp. 87–92, 1999.

[22] M. Yamaguchi, E. Yoshimoto, and S. Kondo, "Pattern regulation in the stripe of zebrafish suggests an underlying dynamic and autonomous mechanism," *Proc. Natl. Acad. Sci. USA.*, vol. 104, no. 12, pp. 4790–4793, 2007.

[23] C. E. Caicedo-Carvajal and T. Shinbrot, "In silico zebrafish pattern formation," *Dev. Biol.*, vol. 315, no. 2, pp. 397–403, 2008.

[24] K. J. Painter, J. M. Bloomfield, J. A. Sherratt, and A. Gerisch, "A Nonlocal Model for Contact Attraction and Repulsion in Heterogeneous Cell Populations," *Bull. Math. Biol.*, vol. 77, no. 6, pp. 1132–1165, 2015.

[25] J. Moreira and A. Deutsch, "Pigment pattern formation in zebrafish during late larval stages: A model based on local interactions," *Dev. Dyn.*, vol. 232, no. 1, pp. 33–42, 2005.

[26] A. Volkening and B. Sandstede, "Iridophores as a source of robustness in zebrafish stripes and variability in Danio patterns," *Nat. Commun.*, vol. 9, no. 1, 2018.

[27] D. M. Parichy and J. M. Turner, "Zebrafish puma mutant decouples pigment pattern and somatic metamorphosis," *Dev. Bio*, vol. 256, no. 2, pp. 242–257, 2003.

[28] J. P. Owen, R. N. Kelsh, and C. A. Yates, "A quantitative modelling approach to zebrafish pigment pattern formation," *eLife*, vol. 9, pp. 1–62, 2020.

[29] E. Gavagnin, J. P. Owen, and C. A. Yates, "Pair correlation functions for identifying spatial correlation in discrete domains," *Phys. Rev. E*, vol. 97, no. 6, pp. 20–22, 2018.

[30] C. B. Kimmel, W. W. Ballard, S. R. Kimmel, B. Ullmann, and T. F. Schilling, "Stages of embryonic development of the zebrafish," *Dev. Dyn.*, vol. 203, no. 3, pp. 253–310, 1995.

[31] D. M. Parichy, M. R. Elizondo, M. G. Mills, T. N. Gordon, and R. E. Engeszer, "Normal table of postembryonic zebrafish development: Staging by externally visible anatomy of the living fish," *Developmental Dynamics*, vol. 238, no. 12, pp. 2975–3015, 2009.

[32] J. W. Moll, A. J. Dziatkowiak, M. Edelman, W. Iljin, E. Ratajczyk-Pakalska, and K. Stengert, "Arterialization of the coronary veins in diffuse coronary arteriosclerosis," *J. Cardiovasc. Surg.*, vol. 16, no. 5, pp. 520–525, 1975.

[33] M. Hirata, K. ichiro Nakamura, T. Kanemaru, Y. Shibata, and S. Kondo, "Pigment cell organization in the hypodermis of zebrafish," *Dev. Dyn.*, vol. 227, no. 4, pp. 497–503, 2003.

[34] C. M. Dooley, A. Mongera, B. Walderich, and C. Nüsslein-Volhard, "On the embryonic origin of adult melanophores: The role of ErbB and kit signalling in establishing melanophore stem cells in zebrafish," *Dev.*, vol. 140, no. 5, pp. 1003–1013, 2013.

[35] A. P. Singh and C. Nüsslein-Volhard, "Zebrafish stripes as a model for vertebrate colour pattern formation," *Curr. Biol.*, vol. 25, no. 2, pp. R81–R92, 2015.

[36] E. Berthier, C. A. Shuman, A. Camerlenghi, L. D. Santis, E. Domack, M. Kirby, M. A. Hemer, D. Roberts, M. Craven, E. Isla, D. J. Demaster, and S. E. Ishman, "Thyroid hormone dependent adult pigment cell lineage," *Science*, vol. 345, no. 6202, pp. 1–5, 2014.

[37] B. Walderich, A. P. Singh, P. Mahalwar, and C. Nüsslein-Volhard, "Homotypic cell competition regulates proliferation and tiling of zebrafish pigment cells during colour pattern formation," *Nat. Commun.*, vol. 7, p. 11462, 2016.

[38] M. Watanabe and S. Kondo, "Comment on "Local reorganization of xanthophores fine-tunes and colors the striped pattern of zebrafish"," *Science*, vol. 348, no. 6232, pp. 1362–4, 2015.

[39] A. P. Singh, U. Schach, and C. Nüsslein-Volhard, "Proliferation, dispersal and patterned aggregation of iridophores in the skin prefigure striped colouration of zebrafish," *Nat. Cell Biol.*, vol. 16, no. 6, pp. 604–611, 2014.

[40] L. Serrano Cardona and E. Muñoz Mata, "Paraninfo Digital," *Early Hum. Dev.*, vol. 83, no. 1, pp. 1–11, 2013.

[41] L. B. Patterson and D. M. Parichy, "Interactions with Iridophores and the Tissue Environment Required for Patterning Melanophores and Xanthophores during Zebrafish Adult Pigment Stripe Formation," *PLoS Genet.*, vol. 9, no. 5, 2013.

[42] H. Yamanaka and S. Kondo, "In vitro analysis suggests that difference in cell movement during direct interaction can generate various pigment patterns in vivo," *Proc. Natl. Acad. Sci. USA.*, vol. 111, no. 5, pp. 1867–1872, 2014.

[43] G. Takahashi and S. Kondo, "Melanophores in the stripes of adult zebrafish do not have the nature to gather, but disperse when they have the space to move," *Pigment Cell and Melanoma Research*, vol. 21, no. 6, pp. 677–686, 2008.

[44] F. J. Van Eeden, M. Granato, U. Schach, M. Brand, M. Furutani-Seiki, P. Haffter, M. Hammerschmidt, C. P. Heisenberg, Y. J. Jiang, D. A. Kane, R. N. Kelsh, M. C. Mullins, J. Odenthal, R. M. Warga, M. L. Allende, E. S. Weinberg, and C. Nüsslein-Volhard, "Mutations affecting somite formation and patterning in the zebrafish, Danio rerio," *Development*, vol. 123, pp. 153–164, 1996.

[45] V. Svetic, G. E. Hollway, S. Elworthy, T. R. Chipperfield, C. Davison, R. J. Adams, J. S. Eisen, P. W. Ingham, P. D. Currie, and R. N. Kelsh, "Sdf1a patterns zebrafish melanophores and links the somite and melanophore pattern defects in choker mutants," *Development*, vol. 134, no. 5, pp. 1011–1022, 2007.

[46] H. G. Frohnhöfer, J. Krauss, H. M. Maischein, and C. Nüsslein-Volhard, "Iridophores and their interactions with other chromatophores are required for stripe formation in zebrafish," *Development (Cambridge)*, vol. 140, no. 14, pp. 2997–3007, 2013.

[47] S. S. Lopes, X. Yang, J. Müller, T. J. Carney, A. R. McAdow, G. J. Rauch, A. S. Jacoby, L. D. Hurst, M. Delfino-Machín, P. Haffter, R. Geisler, S. L. Johnson, A. Ward, and R. N. Kelsh, "Leukocyte tyrosine kinase functions in pigment cell development," *PLoS Genet.*, vol. 4, no. 3, 2008.

[48] A. Fadeev, J. Krauss, A. P. Singh, and C. Nüsslein-Volhard, "Zebrafish Leucocyte tyrosine kinase controls iridophore establishment, proliferation and survival," *Pigment Cell and Melanoma Res.*, vol. 29, no. 3, pp. 284–296, 2016.

[49] A. Fadeev, P. Mendoza-Garcia, U. Irion, J. Guan, K. Pfeifer, S. Wiessner, F. Serluca, A. P. Singh, C. Nüsslein-Volhard, and R. H. Palmer, "ALKALs are in vivo ligands for ALK family receptor tyrosine kinases in the neural crest and derived cells," *Proc. Natl. Acad. Sci. USA.*, vol. 115, no. 4, pp. E630–E638, 2018.

[50] E. S. Mo, Q. Cheng, A. V. Reshetnyak, J. Schlessinger, and S. Nicoli, "Alk and Ltk ligands are essential for iridophore development in zebrafish mediated by the receptor tyrosine kinase Ltk," *Proc. Natl. Acad. Sci. USA.*, vol. 114, no. 45, pp. 12027–12032, 2017.

[51] J. A. Lister, C. P. Robertson, T. Lepage, S. L. Johnson, and D. W. Raible, "Nacre Encodes a Zebrafish Microphthalmia-Related Protein That Regulates Neural-Crest-Derived Pigment Cell Fate," *Development*, vol. 126, no. 17, pp. 3757–3767, 1999.

[52] J. Odenthal, K. Rossnagel, P. Haffter, R. N. Kelsh, E. Vogelsang, M. Brand, F. J. Van Eeden, M. Furutani-Seiki, M. Granato, M. Hammerschmidt, C. P. Heisenberg, Y. J. Jiang, D. A. Kane, M. C. Mullins, and C. Nüsslein-Volhard, "Mutations affecting xanthophore pigmentation in the zebrafish, Danio rerio," *Development*, vol. 123, pp. 391–398, 1996.

[53] J. Krauss, H. G. Frohnhöfer, B. Walderich, H. M. Maischein, C. Weiler, U. Irion, and C. Nüsslein-Volhard, "Endothelin signalling in iridophore development and stripe pattern formation of zebrafish," *Biol. Open*, vol. 3, no. 6, pp. 503–509, 2014.

[54] J. E. Spiewak, E. J. Bain, J. Liu, K. Kou, S. L. Sturiale, L. B. Patterson, P. Diba, J. S. Eisen, I. Braasch, J. Ganz, and D. M. Parichy, "Evolution of Endothelin signaling and diversification of adult pigment pattern in Danio fishes," *PLoS Genet.*, vol. 14, no. 9, pp. 1–21, 2018.

[55] D. M. Parichy, E. M. Mellgren, J. F. Rawls, S. S. Lopes, R. N. Kelsh, and S. L. Johnson, "Mutational analysis of endothelin receptor b1 (rose) during neural crest and pigment pattern development in the zebrafish Danio rerio," *Dev. Biol.*, vol. 227, no. 2, pp. 294–306, 2000.

[56] J. Krauss, P. Astrinidis, H. G. Frohnhöfer, B. Walderich, and C. Nüsslein-Volhard, "Transparent, a gene affecting stripe formation in Zebrafish, encodes the mitochondrial protein Mpv17 that is required for iridophore survival," *Biol. Open*, vol. 2, no. 9, p. 979, 2013.

[57] M. R. Lang, L. B. Patterson, T. N. Gordon, S. L. Johnson, and D. M. Parichy, "Basonuclin-2 requirements for zebrafish adult pigment pattern development and female fertility," *PLoS Genet.*, vol. 5, no. 11, 2009.

[58] S. Fukamachi, M. Sugimoto, H. Mitani, and A. Shima, "Somatolactin selectively regulates proliferation and morphogenesis of neural-crest derived pigment cells in medaka," *Proc. Natl. Acad. Sci. USA.*, vol. 101, no. 29, pp. 10661–10666, 2004.

[59] S. Fukamachi, T. Yada, A. Meyer, and M. Kinoshita, "Effects of constitutive expression of somatolactin alpha on skin pigmentation in medaka," *Gene*, vol. 442, no. 1-2, pp. 81–87, 2009.

[60] F. Van Bebber, A. Hruscha, M. Willem, B. Schmid, and C. Haass, "Loss of Bace2 in zebrafish affects melanocyte migration and is distinct from Bace1 knock out phenotypes," *J. Neurochem.*, vol. 127, no. 4, pp. 471–481, 2013.

[61] P. Mahalwar, A. P. Singh, A. Fadeev, C. Nüsslein-Volhard, and U. Irion, "Heterotypic interactions regulate cell shape and density during color pattern formation in zebrafish," *Biol. Open*, vol. 5, no. 11, pp. 1680–1690, 2016.

[62] D. S. Eom, E. J. Bain, L. B. Patterson, M. E. Grout, and D. M. Parichy, "Long-distance communication by specialized cellular projections during pigment pattern development and evolution," *eLife*, vol. 4, no. DECEMBER2015, pp. 1–25, 2015.

[63] D. S. Eom and D. M. Parichy, "A macrophage relay for long-distance signaling during postembryonic tissue remodeling," *Science*, vol. 355, no. 6331, pp. 1317–1320, 2017.

[64] F. Maderspacher and C. Nüsslein-Volhard, "Formation of the adult pigment pattern in zebrafish requires leopard and obelix dependent cell interactions," *Development*, vol. 130, no. 15, pp. 3447–3457, 2003.

[65] M. Watanabe and S. Kondo, "Changing clothes easily: Connexin41.8 regulates skin pattern variation," *Pigment Cell and Melanoma Res.*, vol. 25, no. 3, pp. 326–330, 2012.

[66] D. S. Eom, S. Inoue, L. B. Patterson, T. N. Gordon, R. Slingwine, S. Kondo, M. Watanabe, and D. M. Parichy, "Melanophore Migration and Survival during Zebrafish Adult Pigment Stripe Development Require the Immunoglobulin Superfamily Adhesion Molecule Igsf11," *PLoS Genet.*, vol. 8, no. 8, 2012.

[67] M. Iwashita, M. Watanabe, M. Ishii, T. Chen, S. L. Johnson, Y. Kurachi, N. Okada, and S. Kondo, "Pigment pattern in jaguar/obelix zebrafish is caused by a Kir7.1 mutation: Implications for the regulation of melanosome movement," *PLoS Genet.*, vol. 2, no. 11, pp. 1861–1870, 2006.

[68] H. Hibino, A. Inanobe, K. Furutani, S. Murakami, I. Findlay, and Y. Kurachi, "Inwardly rectifying potassium channels: Their structure, function, and physiological roles," 2010.

[69] P. Kofuji and E. A. Newman, "Potassium buffering in the central nervous system," *Neuroscience*, vol. 129, no. 4, pp. 1043–1054, 2004.

[70] S. Inoue, S. Kondo, D. M. Parichy, and M. Watanabe, "Tetraspanin 3c requirement for pigment cell interactions and boundary formation in zebrafish adult pigment stripes," *Pigment Cell and Melanoma Res.*, vol. 27, no. 2, pp. 190–200, 2014.

[71] Y. Usui, T. Aramaki, S. Kondo, and M. Watanabe, "The minimal gap-junction network among melanophores and xanthophores required for stripe pattern formation in zebrafish," *Development (Cambridge)*, vol. 146, no. 22, 2019.

[72] S. Kondo and M. Watanabe, "Black, yellow, or silver: Which one leads skin pattern formation?," *Pigment Cell and Melanoma Res.*, vol. 28, no. 1, pp. 2–4, 2015.

[73] I. K. Quigley, J. L. Manuel, R. A. Roberts, R. J. Nuckels, E. R. Herrington, E. L. MacDonald, and D. M. Parichy, "Evolutionary diversification of pigment pattern in Danio fishes: Differential fms dependence and stripe loss in D. albolineatus," *Development*, vol. 132, no. 1, pp. 89–104, 2005.

[74] K. A. Hultman, N. Bahary, L. I. Zon, and S. L. Johnson, "Gene duplication of the zebrafish kit ligand and partitioning of melanocyte development functions to kit ligand a," *PLoS Genetics*, vol. 3, no. 1, pp. 0089–0102, 2007.

[75] M. Watanabe, M. Iwashita, M. Ishii, Y. Kurachi, A. Kawakami, S. Kondo, and N. Okada, "Spot pattern of leopard Danio is caused by mutation in the zebrafish connexin41.8 gene," *EMBO Rep.*, vol. 7, no. 9, pp. 893–897, 2006.

[76] D. A. Lyons, H. M. Pogoda, M. G. Voas, I. G. Woods, B. Diamond, R. Nix, N. Arana, J. Jacobs, and W. S. Talbot, "erbb3 and erbb2 are essential for Schwann cell migration and myelination in zebrafish," *Curr. Biol.*, vol. 15, no. 6, pp. 513–524, 2005.

[77] E. H. Budi, L. B. Patterson, and D. M. Parichy, "Embryonic requirements for ErbB signaling in neural crest development and adult pigment pattern formation," *Development*, vol. 135, no. 15, pp. 2603–2614, 2008.

[78] D. M. Parichy and J. M. Turner, "Temporal and cellular requirements for Fms signaling during zebrafish adult pigment pattern development," *Development*, vol. 130, no. 5, pp. 817–833, 2003.

[79] T. Fukuzawa and H. Ide, "A ventrally localized inhibitor of melanization in Xenopus laevis skin," *Dev. Biol.*, vol. 129, no. 1, pp. 25–36, 1988.

[80] A. Zuasti, J. W. C., P. Samaraweera, and J. T. Bagnara, "Intrinsic PigmentCell Stimulating Activity in the Catfish Integument," *Pigment Cell Res.*, vol. 5, no. 5, pp. 253–262, 1992.

[81] A. Zuasti, "Melanization stimulating factor (MSF) and melanization inhibiting factor (MIF) in the integument of fish," *Microsc. Res. Tech.*, vol. 58, no. 6, pp. 488–495, 2002.

[82] J. Y. Sire, F. Allizard, O. Babiar, J. Bourguignon, and A. Quilhac, "Scale development in zebrafish (Danio rerio)," *J. Anat.*, vol. 190, no. 4, pp. 545–561, 1997.

[83] J. P. Rasmussen, N. T. Vo, and A. Sagasti, "Fish Scales Dictate the Pattern of Adult Skin Innervation and Vascularization," *Dev. Cell*, vol. 46, no. 3, pp. 344–359.e4, 2018.

[84] N. Matsuda, S. Kasagi, T. Nakamaru, R. Masuda, A. Takahashi, and M. Tagawa, "Left-right pigmentation pattern of Japanese flounder corresponds to expression levels of melanocortin receptors (MC1R and MC5R), but not to agouti signaling protein 1 (ASIP1) expression," *Gen. Comp. Endocrinal.*, vol. 262, pp. 90–98, 2018.

[85] M. M. Ollmann and G. S. Barsh, "Down-regulation of melanocortin receptor signaling mediated by the amino terminus of Agouti protein in Xenopus melanophores," *J. Biol. Chem.*, vol. 274, no. 22, pp. 15837–15846, 1999.

[86] R. Cortés, S. Navarro, M. J. Agulleiro, R. Guillot, V. García-Herranz, E. Sánchez, and J. M. Cerdá-Reverter, "Evolution of the melanocortin system," 2014.

[87] L. Cal, P. Suarez-Bregua, P. Comesaña, J. Owen, I. Braasch, R. Kelsh, J. M. Cerdá-Reverter, and J. Rotllant, "Countershading in zebrafish results from an Asip1 controlled dorsoventral gradient of pigment cell differentiation," *Sci. Rep.*, vol. 9, no. 1, p. 3449, 2019.

[88] J. M. Cerdá-Reverter, T. Haitina, H. B. Schiöth, and R. E. Peter, "Gene structure of the goldfish agouti-signaling protein: A putative role in the dorsal-ventral pigment pattern of fish," *Endocrinology*, vol. 146, no. 3, pp. 1597–1610, 2005.

[89] J. B. Gurdon, "A community effect in animal development," *Nature*, vol. 336, no. 6201, pp. 772–774, 1988.

[90] T. Shinbrot, "Simulated morphogenesis of developmental folds due to proliferative pressure," *J. Theor. Biol.*, vol. 242, no. 3, pp. 764–773, 2006.

[91] P. Saunders, "The Chemical Basis of Morphogenesis," *Alan Turing: His Work and Impact*, vol. 237, no. 641, pp. 683–764, 2013.

[92] R. T. Liu, S. S. Liaw, and P. K. Maini, "Two-stage Turing model for generating pigment patterns on the leopard and the jaguar," *Phys. Rev. E - Stat. Nonlinear, and Soft Matter*, vol. 74, no. 1, 2006.

[93] H. Shoji, Y. Iwasa, A. Mochizuki, and S. Kondo, "Directionality of stripes formed by anisotropic reaction-diffusion models," *J. Theor. Biol.*, vol. 214, no. 4, pp. 549–561, 2002.

[94] S. Kondo and R. Asai, "A reaction-diffusion wave on the skin of the marine angelfish Pomacanthus," *Nature*, vol. 376, no. 6543, pp. 765–768, 1995.

[95] H. Shoji, A. Mochizuki, Y. Iwasa, M. Hirata, T. Watanabe, S. Hioki, and S. Kondo, "Origin of directionality in the fish stripe pattern," *Dev. Dyn.*, vol. 226, no. 4, pp. 627–633, 2003.

[96] M. S. Steinberg, "Adhesion in development: An historical overview," *Developmental Biology*, vol. 180, no. 2, pp. 377–388, 1996.

[97] A. Cavagna, A. Cimarelli, I. Giardina, G. Parisi, R. Santagati, F. Stefanini, and M. Viale, "Scale-free correlations in starling flocks," *Proc. Natl. Acad. Sci. USA.*, vol. 107, no. 26, pp. 11865–11870, 2010.

[98] J. E. Green, S. L. Waters, J. P. Whiteley, L. Edelstein-Keshet, K. M. Shakesheff, and H. M. Byrne, "Non-local models for the formation of hepatocyte-stellate cell aggregates," *J. Theor. Biol.*, vol. 267, no. 1, pp. 106–120, 2010.

[99] R. J. Thomas, A. Bennett, B. Thomson, and K. M. Shakesheff, "Hepatic stellate cells on poly(DL-lactic acid) surfaces control the formation of 3D hepatocyte co-culture aggregates in vitro," *Eur. Cells Mater.*, vol. 11, pp. 16–26, 2006.

[100] F. Chalub, Y. Dolak-Struss, P. Markowich, D. Oelz, C. Schmeiser, and A. Soreff, "Model hierarchies for cell aggregation by chemotaxis," *Math. Model. Methods Appl. Sci.*, vol. 16, no. SUPPL. 1, pp. 1173–1197, 2006.

[101] H. G. Othmer and A. Stevens, "Aggregation, blowup, and collapse: The ABC'S of taxis in reinforced random walks," *SIAM J. Appl. Math*, vol. 57, no. 4, pp. 1044–1081, 1997.

[102] H. Murakawa and H. Togashi, "Continuous models for cell-cell adhesion," *J. Theor. Biol.*, vol. 374, pp. 1–12, 2015.

[103] M. J. Simpson, B. J. Binder, P. Haridas, B. K. Wood, K. K. Treloar, D. L. McElwain, and R. E. Baker, "Experimental and Modelling Investigation of Monolayer Development with Clustering," *Bull. Math. Biol.*, vol. 75, no. 5, pp. 871–889, 2013.

[104] E. Hinde, F. Cardarelli, M. A. Digman, and E. Gratton, "In vivo pair correlation analysis of EGFP intranuclear diffusion reveals DNA-dependent molecular flow," *Proc. Natl. Acad. Sci. USA.*, vol. 107, no. 38, pp. 16560–16565, 2010.

[105] B. J. Binder and M. J. Simpson, "Quantifying spatial structure in experimental observations and agent-based simulations using pair-correlation functions," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 88, no. 2, pp. 1–10, 2013.

[106] C. Cutright, Z. Brotherton, L. Alexander, J. Harris, K. Shi, S. Khan, J. Genzer, and S. Menegatti, "Packing density, homogeneity, and regularity: Quantitative correlations between topology and thermoresponsive morphology of PNIPAM-co-PAA microgel coatings," *Appl. Surf. Sci.*, vol. 508, 2020.

[107] N. Topic Popovic, I. Strunjak-Perovic, R. Coz-Rakovac, J. Barisic, M. Jadan, A. Persin Berakovic, and R. Sauerborn Klobucar, "Tricaine methane-sulfonate (MS-222) application in fish anaesthesia," *J. Appl. Ichthyol*, vol. 28, no. 4, pp. 553–564, 2012.

[108] Y. Cakir and S. M. Strauch, "Tricaine (MS-222) is a safe anesthetic compound compared to benzocaine and pentobarbital to induce anesthesia in leopard frogs (Rana pipiens)," *Pharmacol. Reports*, vol. 57, no. 4, pp. 467–474, 2005.

[109] P. J. Rombough, "Ontogenetic changes in the toxicity and efficacy of the anaesthetic MS222 (tricaine methanesulfonate) in zebrafish (Danio rerio) larvae," *Comp. Biochem. Physiol. - A Mol. Integr. Physiol.*, vol. 148, no. 2, pp. 463–469, 2007.

[110] J. W. Lu, Y. J. Ho, Y. J. Yang, H. A. Liao, S. C. Ciou, L. I. Lin, and D. L. Ou, "Zebrafish as a disease model for studying human hepatocellular carcinoma," *J. Gastroenterology*, vol. 21, no. 42, pp. 12042–12058, 2015.

[111] E. Menelaou and K. R. Svoboda, "Secondary motoneurons in juvenile and adult zebrafish: Axonal pathfinding errors caused by embryonic nicotine exposure," *J. Comp. Neurol*, vol. 512, no. 3, pp. 305–322, 2009.

[112] I. Jevtov, T. Samuelsson, G. Yao, A. Amsterdam, and K. Ribbeck, "Zebrafish as a model to study live mucus physiology," *Sci. Rep*, vol. 4, p. 6653, 2014.

[113] A. L. Le Vin, A. Adam, A. Tedder, K. E. Arnold, and B. K. Mable, "Validation of swabs as a non-destructive and relatively non-invasive DNA sampling method in fish," *Mol. Ecol. Resour.*, vol. 11, no. 1, pp. 107–109, 2011.

[114] M. Matthews and Z. M. Varga, "Anesthesia and euthanasia in zebrafish," *ILAR J.*, vol. 53, no. 2, pp. 192–204, 2012.

[115] J. Pirhonen and C. B. Schreck, "Effects of anaesthesia with MS-222, clove oil and CO2 on feed intake and plasma cortisol in steelhead trout (Oncorhynchus mykiss)," *Aquaculture*, vol. 220, no. 1-4, pp. 507–514, 2003.

[116] M. Javadi Moosavi, M. M. Salahi Ardekani, A. Pirbeigi, and S. Ghazi, "The effects of exposure duration to optimal concentration of 2-phenoxyethanol on primary and secondary stress responses in kutum (Rutilus frisii kutum)," *J. Anim. Physiol. Anim. Nutr.*, vol. 99, no. 4, pp. 661–667, 2015.

[117] R. A. Weber, J. J. Pérez-Maceira, J. B. Peleteiro, L. García-Martín, and M. Aldegunde, "Effects of acute exposure to 2-phenoxyethanol, clove oil, MS-222, and metomidate on primary and secondary stress responses in Senegalese sole (Solea senegalensis Kaup 1858)," *Aquaculture*, vol. 321, no. 1-2, pp. 108–112, 2011.

[118] D. Wong, M. A. Von Keyserlingk, J. G. Richards, and D. M. Weary, "Conditioned place avoidance of zebrafish (Danio rerio) to three chemicals used for euthanasia and anaesthesia," *PLoS ONE*, vol. 9, no. 2, p. 88030, 2014.

[119] J. Priborsky and J. Velisek, "A Review of Three Commonly Used Fish Anesthetics," *Rev. Fish. Sci. Aquac.*, vol. 26, no. 4, pp. 417–442, 2018.

[120] J. F. Butterworth and G. R. Strichartz, "Molecular mechanism of local anesthetics: A review," *Anesthesiology*, vol. 72, no. 72, pp. 711–734, 1990.

[121] W. C. Huang, Y. S. Hsieh, I. H. Chen, C. H. Wang, H. W. Chang, C. C. Yang, T. H. Ku, S. R. Yeh, and Y. J. Chuang, "Combined use of MS-222 (Tricaine) and isoflurane extends anesthesia time and minimizes cardiac rhythm side effects in adult zebrafish," *Zebrafish*, vol. 7, no. 3, pp. 297–304, 2010.

[122] T. Martins, E. Diniz, L. M. Félix, and L. Antunes, "Evaluation of anaesthetic protocols for laboratory adult zebrafish (Danio rerio)," *PLoS ONE*, vol. 13, no. 5, pp. 1–12, 2018.

[123] A. M. Valentim, L. M. Félix, L. Carvalho, E. Diniz, and L. M. Antunes, "A new anaesthetic protocol for adult zebrafish (Danio rerio): Propofol combined with lidocaine," *PLoS ONE*, vol. 11, no. 1, pp. 1–12, 2016.

[124] M. Pucéat, D. Garin, and A. Fréminet, "Inhibitory effect of anaesthesia with 2-phenoxyethanol as compared to MS222 on glucose release in isolated hepatocytes from rainbow trout (Salmo gairdneri)," *Comp. Biochem. Physiol. – Part A: Physiol.*, vol. 94, no. 2, pp. 221–224, 1989.

[125] P. Aleström, L. D'Angelo, P. J. Midtlyng, D. F. Schorderet, S. Schulte-Merker, F. Sohm, and S. Warner, "Zebrafish: Housing and husbandry recommendations," *Lab Anim.*, vol. 54, no. 3, pp. 213–224, 2020.

[126] C. J. Clopper and E. S. Pearson, "The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial," *Biometrika*, vol. 26, no. 4, p. 404, 1934.

[127] V. L. Pritchard, J. Lawrence, R. K. Butlin, and J. Krause, "Shoal choice in zebrafish, Danio rerio: The influence of shoal size and activity," *Anim. Behav.*, vol. 62, no. 6, pp. 1085–1088, 2001.

[128] H. GholipourKanani and S. Ahadizadeh, "Use of propofol as an anesthetic and its efficacy on some hematological values of ornamental fish Carassius auratus," *SpringerPlus*, vol. 2, no. 1, pp. 1–5, 2013.

[129] A. M. Shadrin and N. D. Ozerniuk, "Development of the gill system in early ontogenesis of Danio and nine spike stickleback," *Ontogenez*, vol. 33, no. 2, pp. 118–123, 2002.

[130] C. Singleman and N. G. Holtzman, "Growth and maturation in the zebrafish, Danio Rerio: A staging tool for teaching and research," *Zebrafish*, vol. 11, no. 4, pp. 396–406, 2014.

[131] A. Soivio, K. Nyholm, and M. Huhti, "Effects of anaesthesia with MS 222, neutralized MS 222 and benzocaine on the blood constituents of rainbow trout, Salmo gairdneri," *J. Fish. Biol.*, vol. 10, no. 1, pp. 91–101, 1977.

[132] I. M. Cornish and T. W. Moon, "The glucose and lactate kinetics of American eels, Anguilla rostrata (LeSueur), under MS 222 anaesthesia," *J. Fish. Biol.*, vol. 28, no. 1, pp. 1–8, 1986.

[133] L. A. Brown, "Recirculation anaesthesia for laboratory fish," *Lab Anim.*, vol. 21, no. 3, pp. 210–215, 1987.

[134] P. Rombough, "Gills are needed for ionoregulation before they are needed for O2 uptake in developing zebrafish, Danio rerio," *J. Exp. Biol.*, vol. 205, no. 12, pp. 1787–1794, 2002.

[135] D. M. Parichy, "Evolution of danio pigment pattern development," sep 2006.

[136] P. Gomułka, E. Fornal, B. Berecka, A. Szmagara, and E. Ziomek, "Pharmacokinetics of propofol in rainbow trout following bath exposure," *Pol. J. Vet. Sci.*, vol. 18, no. 1, pp. 147–152, 2015.

[137] J. B. Hunn and J. L. Allen, "Movement of Drugs Across the Gills of Fishes," *Annu. Rev. Pharmacol.*, vol. 14, no. 1, pp. 47–54, 1974.

[138] W. E. Morton, "Occupational phenoxyethanol neurotoxicity: A report of three cases," *J. Occup. Med.*, vol. 32, no. 1, pp. 42–45, 1990.

[139] A. M. Zimmer, M. Mandic, K. M. Rourke, and S. F. Perry, "Breathing with fins: Do the pectoral fins of larval fishes play a respiratory role?," *Am J. Physiol. Regul. Integr. Comp. Physiol.*, vol. 318, no. 1, pp. R89–R97, 2020.

[140] J. Brijs, E. Sandblom, M. Axelsson, K. Sundell, H. Sundh, A. Kiessling, C. Berg, and A. Gräns, "Remote physiological monitoring provides unique insights on the cardiovascular performance and stress responses of freely swimming rainbow trout in aquaculture," *Sci. Rep*, vol. 9, no. 1, 2019.

[141] D. A. Smith, S. A. Smith, and S. D. Holladay, "Effect of previous exposure to tricaine methanesulfonate on time to anesthesia in hybrid tilapias," *J. Aquat Anim. Health*, vol. 11, no. 2, pp. 183–186, 1999.

[142] R. J. Chance, G. A. Cameron, M. Fordyce, P. Noguera, T. Wang, C. Collins, C. J. Secombes, and B. Collet, "Effects of repeated anaesthesia on gill and general health of Atlantic salmon, Salmo salar," *Fish Biol.*, vol. 93, no. 6, pp. 1069–1081, 2018.

[143] J. Ortuño, M. A. Esteban, and J. Meseguer, "Effects of four anaesthetics on the innate immune response of gilthead seabream (Sparus aurata L.)," *Fish and Shellfish Immunol.*, vol. 12, no. 1, pp. 49–59, 2002.

[144] B. Y. G. M. Hughes and G. Shelton, "The Mechanism of Gill Ventilation in Three Freshwater Teleosts," *J. Exp. Biol.*, vol. 35, no. 4, pp. 807–823, 1958.

[145] H. Hamada, M. Watanabe, H. E. Lau, T. Nishida, T. Hasegawa, D. M. Parichy, and S. Kondo, "Involvement of Delta/Notch signaling in zebrafish adult pigment stripe patterning," *Dev.*, vol. 141, no. 2, pp. 318–324, 2014.

[146] L. B. Patterson, E. J. Bain, and D. M. Parichy, "Pigment cell interactions and differential xanthophore recruitment underlying zebrafish stripe reiteration and Danio pattern evolution," *Nat Commun*, vol. 5, p. 5299, 2014.

[147] M. Watanabe and S. Kondo, "Is pigment patterning in fish skin determined by the Turing mechanism?," *Trends Genet.*, vol. 31, no. 2, pp. 88–96, 2015.

[148] M. McClure and A. R. McCune, "Evidence for developmental linkage of pigment patterns with body size and shape in Danios (Teleostei: Cyprinidae)," *Evolution*, 2003.

[149] E. J. Crampin, W. W. Hackborn, and P. K. Maini, "Pattern formation in reaction-diffusion models with nonuniform domain growth," *Bull. Math. Biol.*, vol. 64, no. 4, pp. 747–769, 2002.

[150] E. J. Crampin, E. A. Gaffney, and P. K. Maini, "Reaction and diffusion on growing domains: Scenarios for robust pattern formation," *Bull. Math. Biol.*, vol. 61, no. 6, pp. 1093–1120, 1999.

[151] H. Meinhardt and M. Klingler, "A model for pattern formation on the shells of molluscs," *Journal of Theoretical Biology*, 1987.

[152] J. D. Murray, "A Pre-pattern formation mechanism for animal coat markings," *Journal of Theoretical Biology*, 1981.

[153] J. D. Murray and M. R. Myerscough, "Pigmentation pattern formation on snakes," *Journal of Theoretical Biology*, 1991.

[154] R. Plaza, F. Sanchez-Garduazo, P. Padilla, R. Barrio, and P. Maini, "The Effect of Growth and Curvature on Pattern Formation," *J. Dyn. Differ Equations*, vol. 16, no. 4, pp. 1093–1121, 2004.

[155] R. J. Ross, C. A. Yates, and R. E. Baker, "The effect of domain growth on spatial correlations," *Phys A Stat Mech Appl.*, vol. 466, pp. 334–345, jan 2017.

[156] V. A. Kinsler and L. Larue, "The patterns of birthmarks suggest a novel population of melanocyte precursors arising around the time of gastrulation," *Pigment Cell and Melanoma Research*, 2018.

[157] R. A. Spritz, "Genetic Hypomelanoses: Disorders Characterized By Congenital White Spotting - Piebaldism, Waardenburg Syndrome, and Related Genetic Disorders of Melanocyte Development - Clinical Aspects," in *The Pigmentary System: Physiology and Pathophysiology: Second Edition*, 2007.

[158] C. Xiaodong, "Exact stochastic simulation of coupled chemical reactions with delays," *J. Chem. Phys.*, vol. 126, no. 12, pp. 2340–2361, 2007.

[159] M. J. Simpson, K. A. Landman, and B. D. Hughes, "Multi-species simple exclusion processes," *Phys. A Stat Mech Appl.*, vol. 388, no. 4, pp. 399–406, 2009.

[160] J. M. Cerdá-Reverter, M. J. Agulleiro, R. Cortés, E. Sánchez, R. Guillot, E. Leal, B. Fernández-Durán, S. Puchol, and M. Eley, "Involvement of melanocortin receptor accessory proteins (MRAPs) in the function of melanocortin receptors," *Gen. Comp. Endocrinal.*, vol. 188, no. 1, pp. 133–136, 2012.

[161] L. J. Bridge, J. R. King, S. J. Hill, and M. R. Owen, "Mathematical modelling of signalling in a two-ligand G-protein coupled receptor system: Agonist-antagonist competition," *Math. Biosci.*, vol. 223, no. 2, pp. 115–132, 2010.

[162] C. R. Linnen, Y. P. Poh, B. K. Peterson, R. D. Barrett, J. G. Larson, J. D. Jensen, and H. E. Hoekstra, "Adaptive evolution of multiple traits through multiple mutations at a single gene," *Science*, vol. 339, no. 6125, pp. 1312–1316, 2013.

[163] E. M. Wolf Horrell, M. C. Boulanger, and J. A. D'Orazio, "Melanocortin 1 receptor: Structure, function, and regulation," *Front. Genet.*, vol. 7, no. MAY, pp. 1–16, 2016.

[164] E. Sánchez, V. C. Rubio, and J. M. Cerdá-Reverter, "Molecular and pharmacological characterization of the melanocortin type 1 receptor in the sea bass," *Gen. Comp. Endocrinal.*, vol. 165, no. 1, pp. 163–169, 2010.

[165] E. J. Michaud, S. J. Bultman, L. J. Stubbs, and R. P. Woychik, "The embryonic lethality of homozygous lethal yellow mice (A(y)/A(y)) is associated with the disruption of a novel RNA-binding protein," *Genes Dev*, vol. 7, no. 7 A, pp. 1203–1213, 1993.

[166] E. Aberdam, C. Bertolotto, E. V. Sviderskaya, V. De Thillott, T. J. Hemesath, D. E. Fisher, D. C. Bennett, J. P. Ortonne, and R. Ballottit, "Involvement of microphthalmia in the inhibition of melanocyte lineage differentiation and of melanogenesis by agouti signal protein," *J. Biol. Chem.*, vol. 273, no. 31, pp. 19560–19565, 1998.

[167] E. V. Sviderskaya, S. P. Hill, D. Balachandar, G. S. Barsh, and D. C. Bennett, "Agouti signaling protein and other factors modulating differentiation and proliferation of immortal melanoblasts," 2001.

[168] E. Le Pape, T. Passeron, A. Giubellino, J. C. Valencia, R. Wolber, and V. J. Hearing, "Microarray analysis sheds light on the dedifferentiating role of agouti signal protein in murine melanocytes via the Mc1r," *Proc. Natl. Acad. Sci. USA.*, vol. 106, no. 6, pp. 1802–1807, 2009.

[169] M. W. Miller, D. M. Duhl, H. Vrieling, S. P. Cordes, M. M. Ollmann, B. M. Winkes, and G. S. Barsh, "Cloning of the mouse agouti gene predicts a secreted protein ubiquitously expressed in mice carrying the lethal yellow mutation," *Genes Dev*, vol. 7, no. 3, pp. 454–467, 1993.

[170] S. J. Bultman, E. J. Michaud, and R. P. Woychik, "Molecular characterization of the mouse agouti locus," *Cell*, vol. 71, no. 7, pp. 1195–1204, 1992.

[171] H. B. Schiöth, T. Haitina, M. K. Ling, A. Ringholm, R. Fredriksson, J. M. Cerdá-Reverter, and J. Klovins, "Evolutionary conservation of the structural, pharmacological, and genomic characteristics of the melanocortin receptor subtypes," 2005.

[172] R. Guillot, R. Cortés, S. Navarro, M. Mischitelli, V. García-Herranz, E. Sánchez, L. Cal, J. C. Navarro, J. M. Míguez, S. Afanasyev, A. Krasnov, R. D. Cone, J. Rotllant, and J. M. Cerdá-Reverter, "Behind melanocortin antagonist overexpression in the zebrafish brain: A behavioral and transcriptomic approach," *Horm. Behav.*, vol. 82, pp. 87–100, 2016.